hematics

D-3

JUN 27 1932

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

A. B. COBLE
UNIVERSITY OF ILLINOIS

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

G. C. EVANS
RICE INSTITUTE

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
E. T. BELL
W. A. MANNING

HARRY BATEMAN
J. R. KLINE
E. P. LANE

HARRY LEVY
MARSTON MORSE
ALONZO CHURCH

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LIV, Number 3

JULY, 1932

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

Hyperelliptic functions and irrational binary invariants. I. By ARTHUR B. COBLE, ... 425

On the topology of algebroid singularities. By OSCAR ZARISKI, . 453

On a theorem of Eddington. By OSCAR ZARISKI, ... 466

Regular linear systems of curves with the singularities of a given curve as base points. By MARGUERITE LEHR, ... 471

Second note on the celestial sphere. By FRANK MORLEY, ... 489

On the common points of two planar cubics and of two planar cyclides. By FRANK MORLEY and W. K. MORRILL, ... 493

The rectangular five-point. By FRANK MORLEY and R. C. YATES, . 496

The quartic space involutorial transformations with a double conic. By L. A. DYE and F. R. SHARPE, ... 499

A relation between metric and Euclidean spaces. By W. A. WILSON, . 505

On the construction of simple arcs. By GORDON T. WHYBURN, . . 518

Biconnected and related sets. By P. M. SWINGLE, ... 525

A note on spaces having the S property. By GORDON T. WHYBURN, . 536

Remarks on the recurrent behavior of the characteristics on a torus. By AUREL WINTNER, ... 539

Some remarks on the join of two complexes and on invariant subsets of a complex. By EGBERT R. VAN KAMPEN, ... 543

Some additions to the theory of combinators. By H. B. CURRY, . 551

On the overconvergence of sequences of rational functions. By J. L. WALSH, ... 559

On approximation to a mapping function by polynomials. By O. J. FARRELL, ... 571

On infinite orthogonal matrices. By MONROE HARNISH MARTIN, . 579

A note on the convergence of the successive approximations to the solution of an ordinary differential equation. By E. K. HAVILAND, ... 632

THE AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED BY THE J. H. FURST COMPANY
BALTIMORE, MD

HYPERELLIPTIC FUNCTIONS AND IRRATIONAL BINARY INVARIANTS. I.

By Arthur B. Coble.

Introduction. Much of the theory of the theta functions, both in the hyperelliptic and in the general case, is based on the theorem that all functions of given order and characteristic can be expressed linearly in terms of a definite number, k, of such functions. It thus becomes important to have at least one linearly independent set of k functions in terms of which all functions of the given order and characteristic can be linearly expressed. Nevertheless many applications of great interest are based on the existence of quite special three-term linear relations connecting particular functions (cf. [4], p. 441). These special relations must be consequences of the more general linear relations connecting $k+1$ functions. Yet the special relations can be obtained from the more general relations only by some knowledge of the particular behavior of the coefficients in these relations.

Usually the coefficients of the theta relations are values of the theta functions for particular values of the arguments; or, if one prefers, they are the values of given functions with given characteristics for the zero values of the arguments. These coefficients depend only on the moduli of the theta functions. It is therefore desirable, as Wirtinger has pointed out (cf. [5], p. 69), to consider theta relations as relations connecting functions both of the moduli and of the arguments with coefficients in the domain of the ordinary number system.

A similar situation exists in the theory of invariants. Four quadratic covariants of a system of binary forms are necessarily linearly dependent when considered as polynomials in the binary variables alone. Usually however the coefficients of the linear relation are invariants, and in such case the four quadratics are independent in the numerical domain.

The hyperelliptic thetas, defined by a binary form of order $2p+2$, can be exhibited very satisfactorily, in the case $p=2$, by beginning with a numerical domain. It is the object of the present series of papers to consider the possible extensions of the situation for $p=2$ to higher values of p. We outline briefly here the case, $p=2$. Having chosen in a linear space, $S_3(y)$, the five points which determine a coördinate system, and which therefore have only numerical coördinates, the system (∞^4) of quadrics on these five points

maps $S_3(y)$ upon the points of a locus $M_3^3(x)$ in S_4 whose equation may be put into the form, $\sum_{i=1}^{i=6} x_i^3 = 0$ $\{\sum_{i=1}^{i=6} x_i \equiv 0\}$. A binary sextic with roots $t_1, \cdots, t_6$ and differences (ik) has 15 irrational linear invariants (A) of the form $(12)(34)(56)$ of which only five are linearly independent in the domain $R(1)$. Five that are independent map the ∞^3 projectively distinct binary sextics upon the points of $M_3^3(x)$. The sextic also has ten irrational invariants (B) of the second degree and of the form $(12)(13)(23)\cdot(45)(46)(56)$, of which again five are linearly independent. Five which are linearly independent, and which are properly chosen with respect to the five invariants (A), map the ∞^3 binary sextics upon the spaces ξ of an envelope, $M_3^4(\xi)$, of class four, which is the dual form of $M_3^3(x)$ in such wise that x, ξ are point, and tangent space at the point, of $M_3^3(x)$. Thus the moduli of the theta functions are determined algebraically by the choice of point x, or space $S_3(\xi)$, of the "modular manifolds" $M_3^3(x)$, $M_3^4(\xi)$, which themselves have no absolute invariants.

In § 1 we prove, for the general $(2p+2)$-ic, that the system of invariants (A), and the system of invariants (B), have the same dimension, and therefore define modular manifolds, $M_{2p-1}(x)$, $M_{2p-1}(\xi)$ in the same linear space. In § 2 $M_{2p-1}(x)$ is discussed as the map of a linear space, $S_{2p-1}(y)$. The projective construction of a point x of $M_{2p-1}(x)$, determined by a $(2p+2)$-ic with ordered roots, is given in terms of the positions of certain "median" points. For $p=2$, these median points are the ten nodes of $M_3^3(x)$. In §§ 3, 4 the linear relations connecting these median points, and the construction mentioned, are discussed analytically. In § 5 we prove that the x, ξ are dually related, and that x, ξ defined by the same $(2p+2)$-ic are incident.

Returning again to the case $p=2$, when x is chosen on $M_3^3(x)$, the enveloping cone of $M_3^3(x)$ from x is a Kummer cone K whose section by an S_3 not on x is a Kummer surface; and the polar quadric of x as to $M_3^3(x)$ cuts $M_3^3(x)$ in an M_2^6 which is the map from $S_3(y)$ of a Weddle quartic surface. The birational transformation from Weddle to Kummer surface is immediate. If y is a point on the Weddle which maps into x' on M_3^3, the tangent S_3 of M_3^3 at x' is on x, and therefore on K, and contributes, in the section, a tangent plane of the Kummer surface. Hence the domain (∞^5) of the moduli, and of the variables, of the theta functions $p=2$ is represented by pairs of points x, x' on M_3^3 such that the tangent S_3 at x' is on x.

The extension of the Kummer surface to a Kummer manifold K_p of dimension p for the general theta functions of genus p has been developed by Wirtinger. In the hyperelliptic case K_p still exists, but doubtless has many interesting special properties. The author has recently developed the corre-

sponding hyperelliptic extension of the Weddle W_p, and has given the mapping by means of which these manifolds are birationally related. In § 6 some applications to W_3 and to W_p are given. But the question still remains as to whether the K_p and W_p may not be projectively attached to a point of a modular manifold M_{2p-1} in such wise that their birational connection is as evident as in the case $p = 2$.

The author has also proved ([3], pp. 144-148) that the configuration of nodes and tropes of the Kummer, as well as algebraic parametric representations of both the Kummer and the Weddle, can be given in quite simple form in the case $p = 2$ in terms of irrational invariants and covariants of the sextic. We shall examine later the possibility of extending these representations to higher values of p.

As at present planned, part II of this series will be devoted to a more careful examination of the particular case $p = 3$; and part III to a discussion of the loci on $M_{2p-1}(x)$ which correspond to $(2p+2)$-ics with multiple roots of any description.

1. The number of linearly independent linear irrational covariants of weight k for a binary form of order $j = 2k + l$. Let $F = (\alpha t)^j$ be a binary form of order j whose leading coefficient we take for convenience as unity. If the roots of $F = 0$ in nonhomogeneous form are $t = t_1, t_2, \cdots, t_j$, let

$$(ij) = t_i - t_j, \qquad (it) = (t_i - t). \tag{1}$$

If then we select from the j roots k pairs, and form the product of differences,

$$\pi = (12)(34)\cdots(2k-1, 2k)(2k+1, t)\cdots(2k+l, t), \tag{2}$$

this product is an irrational covariant of F of *degree* 1, *weight* k, and *order* l ([1], I, § 3).

The number of distinct covariants of this type is

$$\binom{j}{l}\frac{(2k)!}{k!\,2^k} = \frac{j!}{k!\,l!\,2^k} \qquad (j = 2k + l). \tag{3}$$

However, by using the binary identities, a particular product may be expressed in terms of others with coefficients which are free of the roots, i. e., these covariants are related linearly and with integer coefficients in many ways. We ask for the number,

$$n_{k,l} \qquad (j = 2k + l), \tag{4}$$

which are linearly independent in the ring of integers.

The products π which vanish when $t = t_j$ are those which contain the

factor (jt). The coefficients of (jt) in these products lie in the linear system of irrational linear covariants of weight k and order $l-1$ belonging to the form $F/(jt)$. Thus the number of linearly independent coefficients of (jt) is $n_{k,l-1}$. If π does not contain t_j in the form (jt), π must have the form

$$(1j)(2t)(3t)\cdots(l+1,t)\cdot\pi'.$$

Since $(1j)(2t)=(12)(jt)+(1t)(2j)$, it is clear that, to within products containing (jt), i. e., to within the linear system already accounted for, the terms,

$$\begin{aligned}
&(1j)(2t)(3t)\cdots(l+1,t)\cdot\pi',\\
&(1t)(2j)(3t)\cdots(l+1,t)\cdot\pi',\\
&(1t)(2t)(3j)\cdots(l+1,t)\cdot\pi',\\
&\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&(1t)(2t)(3t)\cdots(l+1,j)\cdot\pi',
\end{aligned}$$

are equal to each other. But the sum of these terms divided by $l+1$ is the polar of t_j as to a covariant of weight $k-1$ and order $l+1$. Since all of these covariants can be expressed in terms of $n_{k-1,l+1}$, we have the recursion formula

$$n_{k,l}=n_{k,l-1}+n_{k-1,l+1} \qquad (l\geq 1). \tag{5}$$

In the derivation of (5) the existence of factors of the type (jt) was assumed. If none such occur, i. e., if we seek $n_{k,0}$ $(j=2k)$, the number of linearly independent invariants of type $(12)(34)\cdots(j-1,j)$, then, as before, there is a linear system which has the factor (12), and of these $n_{k-1,0}$ are linearly independent. The remaining invariants are of type $(13)(24)\pi'$. This type pairs with $(14)(23)\pi'$. The difference of the two contains the factor (12), and is in the linear system accounted for. The sum is twice the polar of t_1, t_2 as to the quadratic $(3t)(4t)\cdot\pi'$ which lies in a linear system containing $n_{k-2,2}$ linearly independent terms. Hence

$$n_{k,0}=n_{k-1,0}+n_{k-2,2}. \tag{6}$$

If we call $r+s$ the *rank* of the number $n_{r,s}$, we observe that the recursion formulae (5), (6) give the value of $n_{k,l}$ in terms of numbers $n'_{k',l'}$ for which the rank $k'+l'$ is either smaller than $k+l$, or equal to $k+l$ when $l'>l$. Since the maximum value of l is j, and then there is a single covariant F, i. e., $n_{0,j}=1$, the formulae (5), (6) suffice to determine the numbers $n_{k,l}$ completely.

A table of early values calculated from these formulae is:

$$n_{0,j} = 1 \qquad (j = 1, 2, \cdots, 10),$$
$$n_{1,l} = 1, 2, 3, \cdots, 9 \qquad (l = 0, 1, 2, \cdots, 8),$$
$$n_{2,l} = 2, 5, 9, 14, 20, 27, 35 \qquad (l = 0, \cdots, 6),$$
$$n_{3,l} = 5, 14, 28, 48, 75 \qquad (l = 0, \cdots, 4),$$
$$n_{4,l} = 14, 42, 90 \qquad (l = 0, 1, 2),$$
$$n_{5,0} = 42.$$

From an inspection of the table we find that

$$(7) \qquad n_{k,l} = \binom{2k+l}{k} - \binom{2k+l}{k-1} = \binom{2k+l}{k-1}\frac{l+1}{k}.$$

It is easy to verify that this value of $n_{k,l}$ satisfies the recursion formulae (5), (6), and therefore it is the value sought.

For the particular case when $l = 0$ and $k = p+1$ the linear covariants become the linear invariants of the binary $(2p+2)$-ic of type (A), i. e., of type $(12)(34)(56)\cdots(2p+1, 2p+2)$. For these it is clear that

(8) *The number of linearly independent irrational linear invariants of the binary* $(2p+2)$*-ic is*

$$\nu = n_{p+1,0} = \binom{2p+2}{p}\frac{1}{p+1} = \binom{2p+2}{p+1}\frac{1}{p+2}.$$

Another type of irrational invariant of the binary $(2p+2)$-ic arises from the norm-curve in S_p with canonical equations,

$$(9) \qquad z_0 = t^p,\ z_1 = t^{p-1},\ z_2 = t^{p-2}, \cdots, z_p = 1.$$

The condition that the $(p+1)$ points of this curve which are determined by $t = t_1, \cdots, t_{p+1}$ be on an S_{p-1} (a condition which entails a degeneration of the norm-curve) is

$$(10) \qquad (1\,2\,3\cdots, p+1) = \begin{vmatrix} t_1^p & t_1^{p-1} \cdots 1 \\ t_2^p & t_2^{p-1} \cdots 1 \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ t_{p+1}^p & t_{p+1}^{p-1} \cdots 1 \end{vmatrix} = (12)(13)(23)\cdots(p, p+1) = 0.$$

The product of two such complementary determinants,

$$(1\,2\,3\cdots, p+1)(p+2, p+3, \cdots, 2p+2),$$

is an irrational invariant of type (B) of the binary $(2p+2)$-ic, whose degree is p. The invariants of type (B) are connected by linear relations with integer coefficients, which arise from determinant identities. According to [1, I, p. 188(70)]

(11) *The number of linearly independent invariants of type* (B) *is the same as the number of linearly independent invariants of type* (A).

In § 5 we develop a duality between these two types of invariants.

2. The modular variety M_{2p-1} in $S_{\nu-1}$. The invariants (A) for two projectively equivalent binary $(2p+2)$-ics are proportional. If then we select ν linearly independent invariants (A) according to § 1 (8), and set

$$(1) \qquad x_i = (i_1 i_2)(i_3 i_4)\cdots(i_{2p+1}\, i_{2p+2}) \qquad (i = 1, \cdots, \nu),$$

the aggregate of projectively distinct and ordered binary $(2p+2)$-ics is mapped upon the points of a variety M_{2p-1} of dimension $2p-1$ in a linear space $S_{\nu-1}$.

The choice of the independent set x_i may be made in many ways. For a particular value of p an available set may be obtained by following out the recurrences of § 1. Thus for $p = 0, 1, 2, 3$ independent sets are

$$\begin{aligned}
&p=0: \ (12);\\
&p=1: \ (12)(34),\ (13)(24);\\
&p=2: \ (12)(34)(56),\ (12)(35)(46),\ (14)(25)(36),\\
&\qquad (13)(24)(56),\ (13)(25)(46);\\
(2)\quad &p=3: \ (12)(34)(56)(78),\ (12)(34)(57)(68),\ (12)(36)(47)(58),\\
&\qquad (13)(24)(56)(78),\ (13)(24)(57)(68),\ (13)(26)(47)(58),\\
&\qquad (12)(35)(46)(78),\ (12)(35)(47)(68),\ (14)(26)(37)(58),\\
&\qquad (13)(25)(46)(78),\ (13)(25)(47)(68),\ (15)(26)(37)(48),\\
&\qquad (14)(25)(36)(78),\ (14)(25)(37)(68).
\end{aligned}$$

A notable property of M_{2p-1} is:

(3) *The modular variety* M_{2p-1} *is a rational variety, which is invariant under a collineation group* $C_{(2p+2)!}$ *which is isomorphic with the symmetric group of order* $(2p+2)!$

For, if the root t_{2p+2} be transformed to $t = \infty$, and at the same time the remaining roots are transformed so that their sum is zero, then these remaining roots, say $y_1, \cdots, y_{2p+1}$, for which

$$(4) \qquad y_1 + y_2 + \cdots + y_{2p+1} = 0,$$

are determined only to within a factor of proportionality. Thus the ordered binary $(2p+2)$-ic is represented by the point y in the linear space $S_{2p-1}(y)$, and conversely a point of $S_{2p-1}(y)$ determines the ordered $(2p+2)$-ic to within a projectivity. In the mapping (1) the binary $(2p+2)$-ic, with roots y_i, ∞ furnishes a point

$$x_i = (y_{i_1} - y_{i_2})(y_{i_3} - y_{i_4}) \cdots, \tag{5}$$

the factor $(i_k i_l)$ which contains t_{2p+2} being simply $+1$ or -1 according as i_k or i_l is $2p+2$. Thus x on M_{2p-1} has coördinates which are rational in the coördinates y of a linear S_{2p-1}, and M_{2p-1} is a rational variety. If the roots of the given $(2p+2)$-ic are permuted, the representative point y is transformed by an operation of a Cremona $G_{(2p+2)!}$ in $S_{2p-1}(y)$—the cross-ratio group of Moore [cf. [1], I, § 7 $(k=1)$]. The varieties $x_i = 0$ in (5) make up the simplest linear system of varieties which is invariant under the Cremona $G_{(2p+2)!}$. When $S_{2p-1}(y)$ is mapped by this linear system upon M_{2p-1}, the Cremona group induces a collineation group $C_{(2p+2)!}$ in $S_{\nu-1}$ under which M_{2p-1} is invariant. The equations of these collineations can be obtained from (1) by permuting the roots t, and expressing the resulting products x'_i linearly in terms of the original products x_i.

The $2p+1$ coördinates y, subject to (4), determine in S_{2p-1} a *base* whose $2p+1$ points have typical coördinates

$$-2p,\ 1,\ 1, \cdots, 1. \tag{6}$$

Clearly all of the p factors $(y_{i_k} - y_{i_l})$ vanish at such a basis point except one at most; all except two at most vanish at two basis points; etc., whence

(7) *M_{2p-1} is the map of $S_{2p-1}(y)$ by the linear system of spreads of order p which contain the $\binom{2p+1}{1}$ points of a basis in S_{2p-1} as $(p-1)$-fold points; and which therefore also contain the $\binom{2p+1}{2}$ basal lines $p-2$ times; the $\binom{2p+1}{3}$ basal planes $p-3$ times; $\cdots$; and finally the $\binom{2p+1}{p-1}$ basal S_{p-2}'s simply.*

For, if we set $z_i = y_i - y_{2p+1}$ $(i = 1, \cdots, 2p)$, then $2p$ of the basis points become the reference points, and the last basis point becomes the unit point. The general spread of order p with $(p-1)$-fold points at the reference points is $\Sigma a_{12\ldots p} z_1 z_2 \cdots z_p = 0$ with $\binom{2p}{p}$ terms. This will also have a $(p-1)$-fold point at the unit point if its polar of order $p-2$ vanishes. This polar has $\binom{2p}{p-2}$ terms of the form $z_1 z_2 \cdots z_{p-2}$, whence the linear system in (7) contains

$$\nu = \binom{2p}{p} - \binom{2p}{p-2} \tag{8}$$

independent members.

We seek now certain sets of points or sets of linear spaces conjugate under the $C_{(2p+2)!}$ in (3), which are of significance for M_{2p-1}. We begin with a set of $\frac{1}{2}\binom{2p+2}{p+1} = \nu(p+2)/2$ *median points.* These are defined as follows: Consider those $(2p+2)$-ics for which the $p+1$ roots, $t_{p+2}, t_{p+3}, \cdots, t_{2p+2}$

become equal to t. Then all of the invariants (A) either vanish or are equal to $\pm (1t)(2t)\cdots(p+1,t)$. Thus the coördinates x_i in (1) become, on dividing by this factor of proportionality, $x_i = \epsilon_i$, where ϵ_i is $+1$, -1, or 0. If, on the other hand, the complementary set of roots all become equal to s, the same invariants (A) vanish as before, and the non-vanishing ones become equal to $\pm (s,p+2)(s,p+3)\cdots(s,2p+2)$, i. e., the coördinates x_i are proportional to the same set ϵ_i as before. Hence

(9) *There exists on M_{2p-1} a conjugate set of $\nu(p+2)/2$ median points $P_{12,\ldots,p+1} = P_{p+2,\ldots,p+2}$ which are the maps of those ordered $(2p+2)$-ics for which one set of $p+1$ roots (or the complementary set) are all equal.*

These points are singular points of the mapping since the ∞^{p-1} projectively distinct $(2p+2)$-ics which depend upon the double ratios of $t_1,\cdots,t_{p+1},t$ all map on the same median point.

Consider now the aggregate of $(2p+2)$-ics for which only p of the roots, say the last p, coincide at t. The invariants (A) which now do not vanish have the form

(10) $$(12)(3t)(4t)\cdots(p+2,t).$$

Of these, according to § 1 (7), there are $n_{1,p} = p+1$ which are linearly independent, and the $(2p+2)$-ics are mapped upon the points of an S_p. As t changes in (10) the mapping point runs over a norm curve of order p in S_p which for $t = t_1, t_2, \cdots, t_{p+2}$ passes through the $p+2$ median points, $P_{1,p+3,\ldots,2p+2}$, $P_{2,p+3,\ldots,2p+2},\cdots, P_{p+2,p+3,\ldots,2p+2}$, in the S_p. For variation of $t_1,\cdots,t_{p+2}$ we obtain the system of ∞^{p-1} norm curves on the base in S_p which consists of these median points. The number of such S_p's is $\binom{2p+2}{p} = (p+1)\nu$. Hence

(11) *There are $\nu(p+1)$ sets of $p+2$ median points whose indices have p common subscripts such that each set lies in an S_p which is contained entirely on M_{2p-1}. The points of such an S_p map binary $(2p+2)$-ics with a p-fold root t. As t varies, the map describes a norm-curve N^p on the $p+2$ median points; and as the remaining roots vary, the N^p runs over a basic system on the $p+2$ points. On each median point there are $2p+2$ such spaces S_p.*

Consider again the aggregate of $(2p+2)$-ics whose last $p-1$ roots coincide at t. The invariants (A) which do not vanish have the form

(12) $$(12)(34)(5t)\cdots(p+3,t).$$

According to § 1 (7), there are $p(p+3)/2$ of these which are linearly in-

dependent. Such $(2p+2)$-ics are mapped upon the points of a manifold M_{p+1} of dimension $p+1$ in a linear space $S_{-1+p(p+3)/2}$. As t varies in (12) the mapping point runs over a norm-curve N^{p-1} (necessarily in an S_{p-1}), and as t takes in succession the values $t_1, \cdots, t_{p+3}$ the mapping point strikes successively a point of one of $p+3$ of the spaces S_p of (11). Since the N^{p-1} is defined by $p+2$ points, the question arises as to whether an S_{p-1} which meets $p+2$ of the spaces S_{p-1} will also meet the remaining one. We find that there are $\infty^{p^2(p+1)/2}$ S_{p-1}'s in $S_{-1+p(p+3)/2}$, and that it is $p(p-1)/2$ conditions that an S_{p-1} shall meet an S_p. Hence there are ∞^p S_{p-1}'s which meet the given $p+2$ S_p's, one on each point of a given S_p. The N^{p-1} on these $p+2$ intersections passes through the point where the S_{p-1} meets the $(p+3)$-th S_p. Hence

(13) *There are* $\binom{2p+2}{p-1}$ *sets of* $(p+3)$ *of the spaces* S_p *of* (11), *each set lying in a linear space* $S_{-1+p(p+3)/2}$. *In* $S_{-1+p(p+3)/2}$ *the* S_{p-1}*'s which meet* $p+2$ *of the* S_p*'s are* ∞^p *in number, one on each generic point of an* S_p. *Such* S_{p-1}*'s also meet the* $(p+3)$*-th* S_p, *and the* $p+3$ *intersections of the* S_{p-1} *with the* $p+3$ S_p*'s lie on a norm-curve* N^{p-1}. *The locus of these* ∞^p N_{p-1}*'s is a manifold* M_{p+1}, *the section of* M_{2p-1} *by* $S_{-1+p(p+3)/2}$, *and the map of binary* $(2p+2)$*-ics with a* $(p-1)$*-fold root.*

From entirely analogous considerations there follows:

(14) *There exists a set of* $\binom{2p+2}{p-k+1}$ *linear spaces* $\{S\}_k$ $(k=1, \cdots, p-1)$ *of dimension* $-1+\binom{p+k+1}{k-1}(p-k+2)/k$ *which cut the modular manifold* M_{2p-1} *each in an* M_{p+k-1}, *the map of projectively distinct binary* $(2p+2)$*-ics with a* $(p-k+1)$*-fold root. A space* $\{S\}_k$ *contains* $(p+k+1)$ *spaces* $\{S\}_{k-1}$. *Through each point of* M_{p+k-1} *there passes a norm-curve* N^{p-k+1} *which meets the* $(p+k+1)$ *spaces* $\{S\}_{k-1}$ *each in one point. The parameters of these* $(p+k+1)$ *points on* N^{p-k+1} *are projective to the remaining distinct roots of the binary* $(2p+2)$*-ic. For the extreme case* $k=p$, *there is one* $\{S\}_p = S_{p-1}$ *containing* M_{2p-1} *and* $\binom{2p+2}{2}$ *spaces* $\{S\}_{p-1}$. *On each point of* M_{2p-1} *there are* $(2p+2)$ *lines each of which crosses* $(2p+1)$ *of the spaces* $\{S\}_{p-1}$.

The theorems given furnish projective constructions for a generic point on M_{2p-1} in terms of the known median points. The median points themselves are obtained from the reference and unit points by harmonic constructions. For example, to obtain the point on M_5 in S_{13} determined by the binary octavic with ordered roots $t_1, \cdots, t_8$ we proceed as follows. In the $\{S\}_1 = S_3$ determined by $t_6 = t_7 = t_8 = t$ we take an N^3 on the 5 S_0's, or median points, determined by $t_1 = t_6 = t_7 = t_8, \cdots, t_5 = t_6 = t_7 = t_8$

with parameters $t_1, \cdots, t_5$ at these five $\{S\}_0$'s in order, and mark the point on this N^3 with parameter t_6. This is the map of the octavic $t_1, \cdots, t_5, t_6, t_6, t_6$. We construct similarly the map of the octavics $t_1, \cdots, t_6, t_5, t_5; \cdots;$ $t_1, \cdots, t_6, t_1, t_1$. These six points lie on an N^2, a conic, with parameters $t_6, t_5, \cdots, t_1$ respectively. On this conic the point t_7 is the map of the octavic $t_1, \cdots, t_7, t_7$. Similarly the maps of the octavics $t_1, \cdots, t_7, t_6; \cdots;$ $t_1, \cdots, t_7, t_1$ are constructed. These seven points lie on a line with parameters $t_7, t_6, \cdots, t_1$ respectively. On this line the point with parameter t_8 is the required point $t_1, \cdots, t_8$.

The manifold M_{2p-1} has, according to (14), $2p+2$ distinct linear rulings. If the M_{2p-1} is mapped as in (7) from an $S_{2p-1}(y)$, then $2p+1$ of the rulings are the maps of the systems of lines each of which is on one of the $2p+1$ basis points in $S_{2p-1}(y)$, and the remaining ruling is the map of the system of norm-curves N^{2p-1} on the basis. These $2p+2$ systems are permuted symmetrically by Moore's $G_{(2p+2)!}$, and thus on M_{2p-1} the $2p+2$ linear rulings are permuted symmetrically by $C_{(2p+2)!}$.

3. Canonical form of the linear identities connecting the invariants (B). The geometric constructions of § 2 have an algebraic parallel which is developed in the next section. In this the median points are assumed as before. These depend upon the division of the $2p+2$ roots into two sets of $p+1$ each, and thus are in one-to-one correspondence with the invariants (B) of § 1 (10), (11), namely:

$$(1) \qquad (B): \quad (i_1 i_2 \cdots i_{p+1})\,(j_1 j_2 \cdots j_{p+1}).$$

These invariants are determinant products, and a canonical choice of sign for each, and a similar choice of a set of independent linear relations connecting them, is useful in the next section.

Let one root, say the last, t_{2p+2}, of the $(2p+2)$-ic be isolated, and let this root occupy the last place in the second determinant in (1). The remaining elements in each of the two determinants are then written in the natural order in their respective determinants so as to produce

$$(2) \qquad d_{l_1 \ldots l_p} = \epsilon(k_1 k_2 \cdots k_{p+1})\,(l_1 l_2 \cdots l_p, 2p+2) \qquad (k_i < k_j \quad \text{and} \quad l_i < l_j \quad \text{if} \quad i < j),$$

where ϵ is $+1$ or -1 according as the permutation

$$k_1 \cdots k_{p+1}\, l_1 \cdots l_p$$

from the natural order is even or odd. Thus the $\binom{2p+1}{p}$ products $d_{l_1 \ldots l_p}$ are determined in sign as soon as the group of subscripts $l_1, \cdots, l_p$ is given.

The linear relations connecting these determinant products are $\binom{2p+2}{p}$ in number. Each relation contains the $p+2$ products which have in one determinant a fixed group of p roots. Thus the relations are of two different types, in the notation introduced above, according as the last root t_{2p+2} is or is not in the fixed group of p roots. For the first type we have

$$(3) \qquad r_{m_1 \dots m_{p-1}} \equiv d_{k_1 m_1 \dots m_{p-1}} + \cdots + d_{k_{p+2}\, m_1 \dots m_{p-1}} = 0.$$

To prove this we observe merely that in the identity as usually written the signs are alternately $+$ and $-$, since each term arises from the preceding one by an inversion. However, this change of sign due to the inversion has been accounted for by ϵ in the definition (2) of $d_{l_1 \dots l_p}$.

In the second type we have a fixed group $m_1, \cdots, m_p$ and a residual group $k_1, \cdots, k_{p+1}, 2p+2$. The terms in the identity are

$$d_{m_1 \dots m_p},\ d_{k_1 \dots k_p},\ d_{k_1 \dots k_{p-1} k_{p+1}}, \cdots,\ d_{k_2 \dots k_{p+1}}.$$

For the same reason as before the signs are like in all terms except perhaps the first. Thus we have to determine ϵ' in

$$\epsilon' d_{m_1 \dots m_p} + d_{k_1 \dots k_p}$$

so that this shall be the pair of terms

$$-(k_1 \cdots k_p\, k_{p+1})(m_1 \cdots m_p, 2p+2) + (m_1 \cdots m_p\ k_{p+1})(k_1 \cdots k_p, 2p+2).$$

Since only the relative signs are to be determined, we may suppose that $k_1 \cdots k_p$ and $m_1, \cdots, m_p$ are already in the natural order. If α transpositions are required to put $k_1 \cdots k_p\, k_{p+1}\, m_1 \cdots m_p$ into the natural order, then $\alpha + p$ transpositions will put $m_1 \cdots m_p\, k_{p+1}\, k_1 \cdots k_p$ into the natural order, since the p transpositions, $(k_1 m_1), \cdots, (k_p m_p)$ change the one into the other. Thus the pair of terms becomes

$$-(-1)^{\alpha} d_{m_1 \dots m_p} + (-1)^{\alpha+p} d_{k_1 \dots k_p},$$

and the second type of linear relation reads as follows:

$$(4) \qquad r_{m_1 \dots m_p} \equiv (-1)^{p+1} d_{m_1 \dots m_p} + d_{k_1 \dots k_p} + d_{k_1 \dots k_{p-1} k_{p+1}} + \cdots + d_{k_2 \dots k_{p+1}} = 0.$$

We now prove that the identities of this second type (4) are obtained from those of the first type (3). Let the roots other than t_{2p+2} be divided into two sets of p and $p+1$ each, say $i_1, \cdots, i_p$ and $j_1, \cdots, j_{p+1}$. Then there are $\binom{p}{0}$ products of type (a_0) or $d_{i_1 \dots i_p}$, $\binom{p}{1}\binom{p+1}{1}$ products of type (a_1) or $d_{i_1 \dots i_{p-1} j_1}$, $\binom{p}{2}\binom{p+1}{2}$ products of type (a_2) or $d_{i_1 \dots i_{p-2} j_1 j_2}, \cdots$, finally $\binom{p}{p}\binom{p+1}{p}$ products of type (a_p) or $d_{j_1 \dots j_p}$. Also there are $\binom{p}{1}$ identities

of type (b_1) or $r_{i_1 \ldots i_{p-1}}$, each containing 1 term of type (a_0) and $p+1$ terms of type (a_1); $\binom{p}{2}\binom{p+1}{1}$ identities of type (b_2) or $r_{i_1 \ldots i_{p-2} j_1}$, each containing 2 terms of type (a_1) and p terms of type (a_2); $\cdots$; finally $\binom{p}{p}\binom{p+1}{p-1}$ identities of type (b_p) or $r_{j_1 \ldots j_{p-1}}$, each containing p terms of type (a_{p-1}) and 2 terms of type (a_p). On adding all identities of the type (b_1); all of the type (b_2); etc., we obtain

$$\begin{aligned} p\,\Sigma\,(a_0) + \Sigma(a_1) &= 0, \\ (p-1)\,\Sigma\,(a_1) + 2\,\Sigma\,(a_2) &= 0, \\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot & \\ 2\,\Sigma\,(a_{p-2}) + (p-1)\,\Sigma\,(a_{p-1}) &= 0, \\ \Sigma\,(a_{p-1}) + p\,\Sigma\,(a_p) &= 0. \end{aligned}$$

The elimination of $\Sigma(a_1), \cdots, \Sigma(a_{p-1})$ from these p equations yields $r_{i_1 \ldots i_p} = (-1)^{p+1}\,\Sigma\,(a_0) + \Sigma\,(a_p) = 0$.

The number of determinant products is $\frac{1}{2}\binom{2p+2}{p+1}$, the number of linear identities of type (3) is $\binom{2p+1}{p-1}$. Since $\frac{1}{2}\binom{2p+2}{p+1} - \binom{2p+1}{p-1} = \binom{2p+2}{p+1} - \binom{2p+2}{p+2}$ is the number of linearly independent determinant products or invariants (B) [cf. § 1 (11)], we have shown that

(5) *With the sign of a determinant product determined as in* (2) *the* $\binom{2p+1}{p}$ *products* $d_{l_1 \ldots l_p}$ *may be arranged as the elements of a p-way determinant* D *of order* $2p+1$ *with lines determined by the ordered subscripts* $l_1, \cdots, l_p$. *An element in two or more lines of the same index is to be zero. Then the* $\binom{2p+1}{p-1}$ *independent linear relations of type* (3) *arise by equating the sum of the elements of any line to zero.*

4. Canonical form of the linear identities connecting the median points; equations of the norm-curves of § 2 in terms of the median points. Setting, as in § 2,

$$(1) \qquad x_i = (i_1 i_2)(i_3 i_4) \cdots (i_{2p+1}\, i_{2p+2}) \qquad (i = 1, \cdots, \nu),$$

let the coördinates ξ_i be contragredient to the point coördinates x in $S_{\nu-1}$. An equation linear in the ξ's is the equation of a point. We prove that the equations of the median points may be affected by such factors of proportionality that their left members satisfy the same system of independent linear relations as the determinant products (B) in § 3 (3).

A particular median point is obtained in (1) when the $p+1$ roots of either of two complementary sets become equal. Let the set containing t_{2p+2} be preferred, and let

$$(2) \qquad \xi_{l_1 \ldots l_p} = 0$$

be the equation of that median point for which $t_{l_1} = t_{l_2} = \cdots = t_{l_p} = t_{2p+2} = t$. If $t_{j_1}, \cdots, t_{j_{p+1}}$ are the remaining roots, the factor $(j_1 t) \cdots (j_{p+1} t)$ is removed from the coördinates x_i in (1), and the residual factors, $+1$, -1, or 0, are the coefficients of the coördinates ξ in (2). For this canonical choice of the factors of proportionality in the coördinates of the median points their equations satisfy the linear identities,

$$\xi_{k_1 m_1 \ldots m_{p-1}} + \cdots + \xi_{k_{p+2} m_1 \ldots m_{p-1}} = 0. \tag{3}$$

We have merely to show that a particular coördinate x_i is either 0 at all of the $p+2$ points in (3), or else is $+1$ at one, -1 at one, and 0 at the p remaining points. If x_i contains a factor $(m_i m_j)$, or a factor $(m_i, 2p+2)$, it vanishes at all of the points of (3), and the identity is satisfied. If no two of the $m_1, \cdots, m_{p-1}, 2p+2$ occur in the same factor of x_i, we may suppose that each occurs last in the particular factor which contains it. This amounts to examining $+x_i$, or $-x_i$ as the case may be. Then x_i or $-x_i$ has the form

$$(k'_1 m_1) \cdots (k'_{p-1} m_{p-1}) (k'_p, 2p+2) (k_r k_s)$$

where $k'_1, \cdots, k'_p$ is some permutation of the $k_1, \cdots, k_{p+2}$ other than k_r and k_s. This vanishes when $t_{m_1}, \cdots, t_{m_{p-1}}, t_{2p+2}$ are all equal to t at every point ξ in (3) except at the two points $\xi_{k_r m_1 \ldots m_{p-1}}$ and $\xi_{k_s m_1 \ldots m_{p-1}}$. At the first of these the factor $(k_r k_s)$ becomes $(t k_s)$, and at the second $(k_r t)$. On removing the factors $(k_s t)$, $(k_r t)$ as part of the factor of proportionality to be deleted, the first point has the value -1, and the second the value $+1$, for the coördinate x_i (or for the coördinate $-x_i$ as the case may be). Hence

(4) *The* $\binom{2p+1}{p}$ *median points* $\xi_{l_1 \ldots l_p}$ *may be arranged as the elements of a p-way determinant* D' *of order* $2p+1$ *with the lines determined by the ordered subscripts* $l_1, \cdots, l_p$. *An element in two or more lines of the same index is to be non-existent. With factors of proportionality chosen as above the* $\binom{2p+1}{p-1}$ *independent linear relations of type* (3) *arise by equating the sum of the elements of any line to zero.*

Let

$$\left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k \qquad (k = 1, \cdots, p+1) \tag{5}$$

denote the polar of the form $(i_{p+2} t) \cdots (i_{p+k+1} t)$ of order k as to the form $(i_1 t) \cdots (i_{p+1} t)$ of order $p+1$ *without however the usual numerical factor;* and let

$$\left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_j \tag{6}$$

denote the value of the polar form (5) for $t = t_j$. We prove the theorem: (7) *The parametric equation of the norm-curve N^{p-k+1} which lies in the space $\{S\}_k$ of § 2 (14) is, in terms of the parameter t, and for $k = 1, \cdots, p$ as follows:*

$$\Sigma_{i_1 \ldots i_{p+k+1}} \left\{ \begin{matrix} i_1 \cdots i_k \\ i_{k+1} \cdots i_{p+k+1} \end{matrix} \right\}^k \xi_{i_1 \ldots i_k i_{p+k+2} \ldots i_{2p+1}} = 0;$$

where Σ is such as to symmetrize for its subscripts. The equation of the point on M_{2p-1} which is the map of the general ordered $(2p+2)$-ic is, in terms of the equations of the median points,

$$\Sigma_{i_1 \ldots i_{2p+1}} \left\{ \begin{matrix} i_1 \cdots i_p \\ i_{p+1} \cdots i_{2p+1} \end{matrix} \right\}^p_{2p+2} \xi_{i_1 \ldots i_p} = 0.$$

The given equations have coefficients which are polars containing t to an order which ensures the existence of the corresponding norm-curve. In order to prove that N^{p-k+1} is the map of the binary $(2p+2)$-ic with ordered distinct roots $t_{i_1}, \cdots, t_{i_{p+k+1}}$, and coincident roots $t_{i_{p+k+2}} = \cdots = t_{i_{2p+2}} = t_{2p+2} = t$, it is necessary only to show that for $t = t_{i_{p+k+1}}$ in N^{p-k+1} we obtain the point $t = t_{i_{p+k+2}}$ on N^{p-k+2} i. e. that

$$(8) \quad \Sigma_{i_1 \ldots i_{p+k+1}} \left\{ \begin{matrix} i_1 \cdots i_k \\ i_{k+1} \cdots i_{p+k+1} \end{matrix} \right\}^k_{i_{p+k+1}} \xi_{i_1 \ldots i_k i_{p+k+2} \ldots i_{2p+1}}$$
$$= (p - k + 3) \Sigma_{i_1 \ldots p+k} \left\{ \begin{matrix} i_1 \cdots i_{k-1} \\ i_k \cdots i_{p+k} \end{matrix} \right\}^{k-1}_{i_{p+k+1}} \xi_{i_1 \ldots i_{k-1} i_{p+k+1} \ldots i_{2p+1}}.$$

For, the norm-curves N in (7) will then be determined by precisely those points which were used in § 2 (14) to define them in succession. In order to prove (8) the following auxiliary formulae are necessary:

$$(9.1) \quad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_{i_{p+k+1}} = (p - k + 2) \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^{k-1}_{i_{p+k+1}};$$

$$(9.2) \quad \left\{ \begin{matrix} i_{p+2} \cdots i_{p+k+1} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^k_{i_1} = - \sum_{r=2}^{r=k+1} \left\{ \begin{matrix} i_{p+2} \cdots i_{p+r-1} \, i_{p+r+1} \cdots i_{p+k+1} \\ i_{p+r} \, i_2 \cdots i_{p+1} \end{matrix} \right\}^{k-1}_{i_1}$$

These give the value of the polar when t is a root, either of the polarizing form, or of the polarized form. In (9.1) the left member is obtained from the polar form $\left\{ \begin{matrix} i_{p+2} \cdots i_{p+k} \\ i_1 \cdots i_{p+1} \end{matrix} \right\}^{k-1}$ of order $p - k + 2$ by operating with $(i_{p+k+1} \, \partial/\partial t)$, and then setting $t = t_{i_{p+k+1}}$. According to Euler's theorem on homogeneous functions, this is the right member. In (9.2) the terms in the polar on the left which contain the factor $(i_1 t)$ vanish for $t = t_{i_1}$. The terms which persist are those which contain i_1 in the form $(i_1 i_{p+r})$ $(r = 2, \cdots, k+1)$. The residual factors are the terms of a polar of next lower order so that the left member can be written as

$$L = \sum_{r=2}^{r=k+1} - (i_{p+r} i_1) \left\{ \begin{matrix} i_{p+2} \cdot \cdot \cdot i_{p+r-1}\, i_{p+r+1} \cdot \cdot \cdot i_{p+k+1} \\ i_2 \cdot \cdot \cdot i_{p+1} \end{matrix} \right\}_{i_1}^{k-1}.$$

On the right the terms of the polars which contain factors of the form $(i_{p+r}\, i_{p+s})$ $(s = 2, \cdot \cdot \cdot, k+1;\ s \neq r)$ will vanish. For, when r takes the value s, the same terms come in with factor $(i_{p+s}\, i_{p+r})$. The only terms which persist when $t = i_1$ are those which contain $(i_{p+r} t)$, and these also reduce to L.

Returning to (8), and isolating the index i_{p+k+1}, the sum on the left breaks up into two parts, namely:

$$\Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdot \cdot \cdot i_{k-1}\, i_{p+k+1} \\ i_k \cdot \cdot \cdot i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k} \xi_{i_1 \ldots i_{k-1}\, i_{p+k+1} \ldots i_{2p+1}}$$
$$+ \Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdot \cdot \cdot i_k \\ i_{k+1} \cdot \cdot \cdot i_{p+k}\, i_{p+k+1} \end{matrix} \right\}_{i_{p+k+1}}^{k} \xi_{i_1 \ldots i_k\, i_{p+k+2} \ldots i_{2p+1}}.$$

On applying (9.1) and (9.2) respectively to these sums, they become

$$(p - k + 2)\, \Sigma_{i_1 \ldots i_{p+k}} \left\{ \begin{matrix} i_1 \cdot \cdot \cdot i_{k-1} \\ i_k \cdot \cdot \cdot i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k-1} \xi_{i_1 \ldots i_{k-1}\, i_{p+k+1} \ldots i_{2p+1}}$$
$$- \Sigma_{i_1 \ldots i_{p+k}} \sum_{j=1}^{j=k} \left\{ \begin{matrix} i_1 \cdot \cdot \cdot i_{j-1}\, i_{j+1} \cdot \cdot \cdot i_k \\ i_j\, i_{k+1} \cdot \cdot \cdot i_{p+k} \end{matrix} \right\}_{i_{p+k+1}}^{k-1} \xi_{i_1 \ldots i_k\, i_{p+k+2} \ldots i_{2p+1}}.$$

Those negative terms in which the lower line of the brace is $i_k \cdot \cdot \cdot i_{p+k}$ will have for coefficients

$$- \sum_{r=k}^{r=p+k} \xi_{i_r i_1 \ldots i_{k-1}\, i_{p+k+2} \ldots i_{2p+1}} = \xi_{i_{p+k+1} i_1 \ldots i_{k-1}\, i_{p+k+2} \ldots i_{2p+1}} \quad [\text{cf. } (3)].$$

Thus these negative terms combine with the positive terms to raise the coefficient $(p - k + 2)$ to $(p - k + 3)$ as on the right of (8).

5. Duality of the invariants (A) with respect to the invariants (B). The final equation of § 4 (7) which expresses the general point of the modular manifold M_{2p-1} in terms of the median points contains all of the median points. The coefficients in it, though formed for isolated t_{2p+2}, are in reality symmetric in all of the roots. The equation might be written in the more symmetric form

$$(1) \qquad \Sigma \left\{ \begin{matrix} i_{p+2} \cdot \cdot \cdot i_{2p+2} \\ i_1 \cdot \cdot \cdot i_{p+1} \end{matrix} \right\}^{p+1} \xi_{i_1 \ldots i_{p+1};\ i_{p+2} \ldots i_{2p+2}} = 0.$$

The earlier form has, however, the advantage in that the summation, and the signs of the terms, are more easily defined and we retain it. The coefficients are the apolarity invariants of two complementary factors of order $p+1$ of the binary $(2p+2)$-ic. Each term of such an invariant is itself an invariant (A). In fact

$$(1.1) \qquad \left\{ \begin{matrix} i_{p+2} \cdot \cdot \cdot i_{2p+2} \\ i_1 \cdot \cdot \cdot i_{p+1} \end{matrix} \right\}^{p+1} = \Sigma\, (i_1 i_{p+2})(i_2 i_{p+3}) \cdot \cdot \cdot (i_{p+1} i_{2p+2})$$

where Σ refers to the $(p+1)!$ permutations of $i_1, \cdot \cdot \cdot, i_{p+1}$, or of $i_{p+2}, \cdot \cdot \cdot, i_{2p+2}$. We make then a linear transformation from the coördinates x of § 2 (1) to coördinates defined by these apolarity invariants. Again we fix i_{2p+2} to be $2p+2$ and set

$$(2) \qquad x_{i_{p+2} \ldots i_{2p+1}} = \left\{ \begin{matrix} i_{p+2} \cdot \cdot \cdot i_{2p+1}, 2p+2 \\ i_1 \cdot \cdot \cdot i_{p+1} \end{matrix} \right\}^{p+1}.$$

With respect to these new coördinates or new invariants (A) we first observe that they satisfy the same set of independent linear relations as the determinant products, or invariants (B) in § 3 (3), or as the median points in § 4 (3), namely:

$$(3) \qquad r_{m_1 \ldots m_{p-1}} \equiv x_{k_1 m_1 \ldots m_{p-1}} + \cdot \cdot \cdot + x_{k_{p+2} m_1 \ldots m_{p-1}} = 0.$$

For, the identity

$$(4) \qquad \left\{ \begin{matrix} k_1 \\ k_2 \cdot \cdot \cdot k_{p+2} \end{matrix} \right\}^1 + \left\{ \begin{matrix} k_2 \\ k_1 k_3 \cdot \cdot \cdot k_{p+2} \end{matrix} \right\}^1 + \cdot \cdot \cdot + \left\{ \begin{matrix} k_{p+2} \\ k_1 \cdot \cdot \cdot k_{p+1} \end{matrix} \right\}^1 \equiv 0,$$

is verified by noting that the term $(k_2 k_1)(k_3 t) \cdot \cdot \cdot (k_{p+2} t)$ of the first of the polars in (4) is cancelled by the term $(k_1 k_2)(k_3 t) \cdot \cdot \cdot (k_{p+2} t)$ of the second, etc. If this identity is polarized further with respect to

$$(m_1 t) \cdot \cdot \cdot (m_{p-1} t)(2p+2, t),$$

the identity (3) is obtained. These relations ensure that the number of linearly independent coördinates x in (2) is not greater than the number, ν, of linearly independent invariants. We ensure that this number is as great as ν by expressing an individual invariant (A) in terms of the invariants x in (2) by the formula,

$$(5) \qquad \frac{(p+2)!}{2} \cdot (12)(34) \cdot \cdot \cdot (2p+1, 2p+2) = \Sigma\, (-1)^p\, x_{135 \ldots, 2p-1},$$

where the 2^p terms in Σ arise from the one given by the operations of the G_{2^p} generated by the p transpositions, $(12), (34), \cdot \cdot \cdot, (2p-1, 2p)$, provided that, when a transposition is applied to a term, the sign of the term is also to be changed. Thus, for $p = 2$, and $p = 3$, we have

$$12 \cdot (12)(34)(56) = \quad x_{13} - x_{23} - x_{14} + x_{24},$$
$$60 \cdot (12)(34)(56)(78) = -x_{135} + x_{235} + x_{145} - x_{245} + x_{136} - x_{236} - x_{146} + x_{246}.$$

To prove (5) we note first that the transposition (12) changes the sign

on the right, whence Σ contains the factor (12); and, by the same argument, Σ contains the factors $(34), \cdots, (2p-1, 2p)$. As a function of the differences of the roots linear in each root Σ must then also contain the factor $(2p+1, 2p+2)$. In order to verify however that Σ does not vanish identically, and at the same time to check the numerical factor in (5), let the odd-numbered roots all be equal to r and the even numbered roots all be equal to s. Then the left member of (5) has the value $(rs)^p \cdot (p+2)!/2$. Each term on the right has the form

$$(-1)^k \left\{ \begin{matrix} r\,r \cdots r\; s\,s \cdots s \\ s\,s \cdots s\; r\,r \cdots r \end{matrix} \right\}^{p+1}$$

where k is the number of r's in the upper line and $k = 0, 1, \cdots, p$ since $t_{2p+2} = s$ always. This value will however arise from $\binom{p}{k}$ terms since in general the $(p-k)$ roots s in the upper line other than the last may be any of $t_2, t_4, \cdots, t_{2p}$. In forming the apolarity invariant the roots s, r in the lower line are distributed in any order with respect to the roots r, s in the upper line so that the non-vanishing terms contribute

$$k!\,(sr)^k \cdot (p+1-k)!\,(rs)^{p+1-k} = (-1)^k\,k!\,(p+1-k)!\,(rs)^{p+1}.$$

The value of Σ is therefore $[\sum_{k=0}^{k=p} \binom{p}{k}\, k!\,(p+1-k)!\,](rs)^{p+1}$. Since $\binom{p}{k} = p!/k!\,(p-k)!$ this can be written as

$$p!\,(rs)^{p+1}\,[\sum_{k=0}^{k=p} (p+1-k)] = p!\,(rs)^{p+1} \cdot (p+1)(p+2)/2$$

which coincides with the left member.

For values of p beyond three no convenient choice of ν independent variables $x_{m_1 \ldots m_p}$, nor of independent combinations of them, seems to exist. Even for $p = 2, 3$ the best procedure is to select a set of $\nu + 1$ combinations whose sum is identically zero (for $p = 2$ cf. [1], I, p. 167; for $p = 3$, cf. Part II of this account). Thus, in the general case, it seems best to retain all the variable point coördinates $x_{m_1 \ldots m_p}$, and therefore to carry along the linear relations (3) on these coördinates. The question then arises as to the nature of the dual coördinates ξ in $S_{\nu-1}$. If the variables were independent we should have dual coördinates x, ξ with the incidence condition $\Sigma\, x_i \xi_i = 0$. When the linear relations are introduced, the coefficients ξ of this condition can be modified into

$$\Sigma_i x_i \xi_i + \Sigma \lambda_{m_1 \ldots m_{p-1}}\, r_{m_1 \ldots m_{p-1}} = 0.$$

If we write this as $\Sigma\, x_i \xi'_i$, the values of the ξ'_i are

$$\xi'_{m_1 \ldots m_p} = \xi_{m_1 \ldots m_p} + \lambda_{m_2 \ldots m_p} + \lambda_{m_1 m_3 \ldots m_p} + \cdots + \lambda_{m_1 m_2 \ldots m_{p-1}}, \tag{6}$$

where the constants λ are at our disposal. It is natural to attempt to choose these $\binom{2p+1}{p-1}$ constants λ so that the coördinates ξ' satisfy the same system of $\binom{2p+1}{p-1}$ linear relations as the coördinates x in (3). If

$$(7) \qquad r_{m_2 \dots m_p} \equiv \xi'_{k_1 m_2 \dots m_p} + \xi'_{k_2 m_2 \dots m_p} + \cdots + \xi'_{k_{p+2} m_2 \dots m_p} = 0,$$

then from (6) there follows that

$$\begin{aligned}(8) \quad & (p+2)\lambda_{m_2 \dots m_p} + \lambda_{k_1 m_3 \dots m_p} + \lambda_{k_1 m_2 m_4 \dots m_p} + \cdots + \lambda_{k_1 m_2 m_3 \dots m_{p-1}} \\ & + \lambda_{k_2 m_3 \dots m_p} + \lambda_{k_2 m_2 m_4 \dots m_p} + \cdots + \lambda_{k_2 m_2 m_3 \dots m_{p-1}} \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & + \lambda_{k_{p+2} m_3 \dots m_p} + \lambda_{k_{p+2} m_2 m_4 \dots m_p} + \cdots + \lambda_{k_{p+2} m_2 m_3 \dots m_{p-1}} \\ = & -\xi_{k_1 m_2 \dots m_p} - \xi_{k_2 m_2 \dots m_p} - \cdots - \xi_{k_{p+2} m_2 \dots m_p}.\end{aligned}$$

If the $\binom{2p+1}{p-1}$ equations (8) in the $\binom{2p+1}{p-1}$ constants λ can be solved for the λ's in terms of the ξ's, the relations (7) can be satisfied. That the solution is possible (i. e., that the determinant of the system is not zero) can be proved by showing that any individual $\lambda_{m_2 \dots m_p}$ can be obtained by taking proper linear combinations of the left members of equations (8). For, if that equation derived from (8) by the interchange of m_2 and k_1 be written, the difference of the two left members is

$$\begin{aligned}(9) \quad (\lambda_{m_2} - \lambda_{k_1}) \{ & (p+1)\lambda_{m_3 \dots m_p} + \lambda_{k_2 m_4 \dots m_p} + \cdots + \lambda_{k_2 m_3 \dots m_{p-1}} \\ & + \cdots\cdots\cdots\cdots\cdots \\ & + \lambda_{k_{p+2} m_4 \dots m_p} + \cdots + \lambda_{k_{p+2} m_3 \dots m_{p-1}} \}.\end{aligned}$$

In (9) the multiplication is symbolic, i. e., $\lambda_{m_2} \lambda_{m_3} \cdots \lambda_{m_p} = \lambda_{m_3 \dots m_p}$. The part of (9) within the brace is the left member of (8) formed for $p' = p - 1$. Assuming that this latter system can be solved for $\lambda_{m_3 \dots m_p}$, the original system can be solved for $\lambda_{m_2 m_3 \dots m_p} - \lambda_{k_1 m_3 \dots m_p}$, i. e., for the difference of $\lambda_{m_2 \dots m_p}$ and any λ which has $p-2$ subscripts in common with it. On adding such properly chosen differences to (8) the resulting left member is a multiple of $\lambda_{m_2 \dots m_p}$. Thus the system (8) can be solved for the λ's if it can be solved in the earliest case $p = 2$. For this case the five left members have the typical form,

$$4\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5$$

and the possibility of the solution is obvious.

At the moment we do not need the explicit values of the new coördinates ξ' in (6); but they are useful later, and therefore we give the solution of the system (8) and the resulting values of the ξ''s in (6). So long as only linear expressions are in question, it is legitimate to use the symbolic multiplication,

$$\xi_{123\ldots} = \xi_1\xi_2\xi_3\cdots = \xi_1\xi_{23}\cdots = \cdots;$$
$$\lambda_{123\ldots} = \lambda_1\lambda_2\lambda_3\cdots = \lambda_1\lambda_{23}\cdots = \cdots. \tag{10}$$

If the $2p+1$ subscripts are divided into complementary sets $m_1, \cdots, m_{p-1}$ and $k_1, \cdots, k_{p+2}$, and if Σ, S refer respectively to sums which symmetrize with respect to these sets, the equation (8) can be written in the form

$$\frac{(p+2)!}{2}\,[(p+2)\lambda_{m_1\ldots m_{p-1}} + \Sigma\lambda_{m_1\ldots m_{p-2}}S\lambda_{k_1}] = -\frac{(p+2)!}{2}\,\xi_{m_1\ldots m_{p-1}}S\xi_{k_1}. \tag{11}$$

Due to the symmetry of the system (8) in all the indices we may assume that the solution whose existence was proved above has the form,

$$\frac{(p+2)!}{2}\,\lambda_{m_1\ldots m_{p-1}} = \sum_{j=0}^{p-1} a_j\Sigma\,\xi_{m_1\ldots m_{p-1-j}}S\xi_{k_1\ldots k_{j+1}}. \tag{12}$$

If we substitute this value in (11), and note that the coefficient of $\xi_{m_1\ldots m_{p-1-j}}\xi_{k_1\ldots k_{j+1}}$ on the right is zero, except that for $j=0$ it is $-(p+2)!/2$, then $a_0, \cdots, a_{p-1}$ are determined by

$$\begin{aligned} &(2p+1)a_0 + (p-1)(p+1)a_1 = -(p+2)/2, \\ &j(j+1)a_{j-1} + (j+1)(2p-2j+1)a_j \\ &\qquad + (p-j-1)(p-j+1)a_{j+1} = 0 \\ &[a_p = 0,\ j = 1, \cdots, p-1]. \end{aligned} \tag{13}$$

An easy verification shows that

$$a_0 = -(p+1)!/2, \quad a_j = (-1)^{j-1}\,j!\,(p-j+1)!/2$$

or

$$\frac{(p+2)!}{2}\,\lambda_{m_1\ldots m_{p-1}} = -\frac{(p+1)!}{2}\,\xi_{m_1\ldots m_{p-1}}S\xi_{k_1} + \sum_{j=1}^{p-1}\frac{(-1)^{j-1}\,j!\,(p-j+1)!}{2}\,\Sigma\xi_{m_1\ldots m_{p-1-j}}S\xi_{k_1\ldots k_{j+1}}. \tag{14}$$

Let the equation (6) be written in the form,

$$\xi'_{m_1\ldots m_p} = \xi_{m_1\ldots m_p} + \mu_{m_1\ldots m_p}. \tag{15}$$

Then

$$\frac{(p+2)!}{2}\,\mu_{m_1\ldots m_p} = \sum_{j=0}^{p}\alpha_j\ \Sigma'\xi_{m_1\ldots m_{p-j}}S'\xi_{k_1\ldots k_j}, \tag{16}$$

where Σ', S' refer to complementary sets $m_1, \cdots, m_p$; $k_1, \cdots, k_{p+1}$. On substituting the values of λ as given in (14) in the value of μ obtained from (6), we find that

$$\alpha_j = ja_{j-1} + (p-j)a_j \qquad (j = 0, \cdots, p).$$

Using the values of the a_j in (14) this yields

$$\alpha_0 = -\frac{(p+1)!\,p}{2}, \quad \alpha_j = (-1)^j (p-j+1)!\,j! \qquad (j=1,\cdots,p),$$

whence

$$(17) \qquad \frac{(p+2)!}{2}\,\xi'_{m_1\ldots m_p} = (p+1)!\,\xi_{m_1\ldots m_p} + \sum_{j=1}^{p} (-1)^j\, j!\,(p-j+1)!\,\Sigma' \xi_{m_1\ldots m_{p-j}}\, S'\xi_{k_1\ldots k_j}.$$

We prove now that

$$(18) \quad \textit{With } x_{m_1\ldots m_p} = \left\{ \begin{matrix} m_1\cdots m_p, 2p+2 \\ k_1\cdots k_{p+1} \end{matrix} \right\}^{p+1} \text{[cf. (2)], and}$$

$$\xi_{m_1\ldots mp} = \epsilon(k_1\cdots k_{p+1})(m_1\cdots m_p, 2p+2) \quad \text{[cf. § 3 (2)],}$$

the aggregate of projectively distinct and ordered binary $(2p+2)$-ics is mapped upon the points x of a manifold $M_{2p-1}(x)$ in $S_{\nu-1}$, and also upon the spaces ξ of a manifold $M_{2p-1}(\xi)$ in $S_{\nu-1}$ which is dually related to $M_{2p-1}(x)$ in such wise that the SAME *ordered $(2p+2)$-ic determines* INCIDENT *point x and space ξ. Under permutation of the roots of the $(2p+2)$-ic the point x and space ξ are subject to the operations of the* SAME *collineation group, $C_{(2p+2)!}$, for which also the manifolds M are invariant.*

For, the variables ξ, defined as in (18), satisfy according to § 3 (3) the same set of linear relations as the variables x [cf. (3)]; and therefore, as proved above, they may be taken as dual coördinates with the incidence condition

$$(19) \quad \Sigma x_{m_1\ldots m_p}\,\xi_{m_1\ldots m_p} = 0 \quad (m_1,\cdots,m_p = 1,\cdots,2p+1;\ m_i \neq m_j).$$

The x's are of degree $(p+1)$ in the differences of the roots, the ξ's of degree $p(p+1)$, so that $x_i\xi_i$ is of degree $(p+1)^2$. On interchange of t_{mi}, t_{mj} the ξ's are permuted with a change of sign throughout, while the x's are merely permuted. Hence $(x\xi)$ must contain $(m_i m_j)$ as a factor, and thus must contain as a factor the $p(2p+1)$ differences which is impossible since its degree is $(p+1)^2$. Thus $(x\xi)$, as a function of the roots, must vanish. When the roots are permuted, the variables x, ξ are permuted in the same way, and the two permutations are dual forms of the same collineation in $S_{\nu-1}$.

When the $M_{2p-1}(x)$ is regarded as the map of the linear space $S_{2p-1}(y)$ as in § 2 (7), the ordered binary $(2p+2)$-ic determines a point y in S_{2p-1} and also the map x of this point y on $M_{2p-1}(x)$ in $S_{\nu-1}$; but it determines also as in (18) a linear space ξ in $S_{\nu-1}$, and therefore a particular member of the linear mapping system in $S_{2p-1}(y)$. We determine this particular

member in the following way and obtain thereby a new interpretation of the incidence relation.

With basal coördinates y in $S_{2p-1}(y)$ such that

$$y_1 + y_2 + \cdots + y_{2p+1} = 0, \tag{20}$$

and $2p+1$ basis points of which the i-th is

$$y_i = -2p, \qquad y_j = 1 \qquad (j \neq i), \tag{21}$$

the equation in terms of the parameter t of the norm-curve N^{2p-1}, which passes through the base points with parameters $t_1, \cdots, t_{2p+1}$ respectively, is

$$y_i - y_j = (ij)\Pi/(it)(jt), \qquad \Pi = (1t)(2t)\cdots(2p+1, t); \tag{22}$$

or

$$2py_i = \Sigma_j (ij)\Pi/(it)(jt).$$

This is in effect the space $\{S\}_1 = S_{2p-1}(y)$ of § 4 (7) when the root t_{2p+2} is transformed to ∞ and the other roots are transformed into $y_1, \cdots, y_{2p+1}$.

For dual coördinates in $S_{2p-1}(y)$ we use $\eta_1, \cdots, \eta_{2p+1}$ where

$$\eta_1 + \eta_2 + \cdots + \eta_{2p+1} = 0; \tag{23}$$

and

$$\eta_1 y_1 + \eta_2 y_2 + \cdots + \eta_{2p+1} y_{2p+1} = 0. \tag{24}$$

is the incidence condition of point y and space η. The norm curve N^{2p-1} is a locus of spaces η as well as a locus of points y. The dual parametric equation of N^{2p-1} is

$$\eta_i = (-1)^{i+1}(123\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} \qquad (i = 1, \cdots, 2p-1). \tag{25}$$

For, we observe first of all that the linear relation (23) is satisfied due to the linear relation,

$$\sum_{i=1}^{2p+1} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} = 0, \tag{26}$$

among $2p+1$ perfect $(2p-1)$-th powers of linear forms. Secondly we prove that the space $\eta(t)$ in (25) will cut the curve $y(t')$ in (22) in $2p-1$ points determined by $(t't)^{2p-1} = 0$. Indeed the incidence condition (24) can be modified, by using (20) and then (22), to read

$$\begin{aligned}
&\sum_{i=1}^{2p} \eta_i [y_i(t') - y_{2p+1}(t')] \\
&= \sum_{i=1}^{2p} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p+1)(it)^{2p-1} \\
&\qquad \times (i, 2p+1)\Pi(t')/(it')\,(2p+1, t') \\
&= \prod_{i=1}^{2p} (i, 2p+1) \cdot \sum_{i=1}^{2p} (-1)^{i+1}(12\cdots, i-1, i+1, \cdots, 2p, t') \cdot (it)^{2p-1} = 0.
\end{aligned}$$

This reduces, by using the relation (26) in which t_{2p+1} is replaced by t', to $-[\prod_{i=1}^{2p}(i, 2p+1)]\cdot(123\cdots 2p)\cdot(t't)^{2p-1}=0$. Thus the equation $(\eta y)=0$ of the space $\eta(t)$ of the N^{2p-1} in (25) is

$$(27)\quad y_1(234\cdots,2p+1)(1t)^{2p-1}-y_2(134\cdots,2p+1)(2t)^{2p-1}+\cdots+y_{2p+1}(123\cdots,2p)(2p+1,t)^{2p-1}=0.$$

For given t, (27) is the equation in variables y of the space η which has $(2p-1)$-point contact with N^{2p-1} at the point t; for given y, (27) is the equation in variable t which determines the $2p-1$ parameters t of spaces η of N^{2p-1} on the point y.

A binary form of odd order $2p-1$ such as (27) has a covariant of degree p, order p, and weight $p(p-1)$, namely, its *canonizant*, such that the form can be expressed as a sum of the $(2p-1)$-th powers of the p linear factors of the canonizant. If the form can be expressed as a sum of less than p such powers the canonizant vanishes identically. If then the form is

$$k_1(1t)^{2p-1}+k_2(2t)^{2p-1}+\cdots+k_p(pt)^{2p-1},$$

the canonizant is, to within a numerical factor,

$$k_1k_2\cdots k_p(12\cdots p)^2(1t)(2t)\cdots(pt).$$

Hence the canonizant of (27) is made up of a sum of such products, each product arising from p terms of (27). We examine the above typical product when $k_1=(23\cdots,2p+1)y_1$, etc., and observe first that it contains $[(\Delta_{2p+1})^{\frac{1}{2}}]^{p-1}$ as a factor where $(\Delta_{2p+1})^{\frac{1}{2}}$ is the product of differences, $(12\cdots,2p+1)$. The residual factor is

$$y_1y_2\cdots y_p(12\cdots pt)(p+1,\cdots,2p+1)=d_{12\ldots p}\,y_1y_2\cdots y_p.$$

There is in addition a sign depending on the number of minus signs occurring in the p terms in (27). This sign will change as the p terms selected from (27) change, but the change is precisely that accounted for in the definition of $d_{12\ldots p}$ in § 3 (2). Hence

(28) *The canonizant of the binary* $(2p-1)$*-ic in* (27) *is, to within a numerical factor and the factor,* $[(\Delta_{2p+1})^{\frac{1}{2}}]^{p-1}$,

$$\Sigma_{12\ldots p}\,d_{12\ldots p}\,y_1y_2\cdots y_p.$$

This expression for the canonizant can be modified so as to take the form (19) of the incidence condition in $S_{\nu-1}$. For, the $d_{12\ldots p}$ satisfy according to § 3 (3) the same system of linear relations as the coördinates

$x_{12\dots p}$ in (3), and therefore their coefficients may be modified as in (17) the $\xi_{12\dots p}$ were modified into $\xi'_{12\dots p}$. Thus the canonizant (28) can be written as

$$\frac{(p+2)!}{2}\,\Sigma_{12\dots p}\,d_{12\dots p}\,y_1y_2\cdots y_p$$
$$=\Sigma_{12\dots p}\,d_{12\dots p}\{(p+1)!\,y_1y_2\cdots y_p$$
$$+\sum_{j=1}^{p}(-1)^j\,j!\,(p-j+1)!\,\Sigma''y_1\cdots y_{p-j}\,S''y_{p+1}\cdots y_{p+j}\},$$

where Σ'' refers to a symmetrization as to $1,\cdots,p$, and S'' as to $p+1,\cdots,2p+1$. The coefficient of $d_{12\dots p}$ is now

$$\left\{\begin{matrix} y_1\cdots y_p\,\infty \\ y_{p+1}\cdots y_{2p+1}\end{matrix}\right\}^{p+1}. \tag{29}$$

For, the value of this polar is

$$S''(y_1-y_{p+1})(y_2-y_{p+2})\cdots(y_p-y_{2p})$$

i.e. it is a sum of $(p+1)!$ products. In each product there is a term $y_1y_2\cdots y_p$. Moreover in the sum there are $(p+1)!\binom{p}{j}$ terms of the type $(-1)^j\,y_1\cdots y_{p-j}\,y_{p+1}\cdots y_{p+j}$, and $\Sigma''y_1\cdots y_{p-j}\,S''y_{p+1}\cdots y_{p+j}$ contains $\binom{p}{j}\binom{p+1}{j}$ terms of this type, whence each term arises $(p-j+1)!\,j!$ times. Hence

(30) *To within an additional numerical factor the canonizant* (28) *can be written as*

$$\Sigma_{12\dots p}\,d_{12\dots p}\left\{\begin{matrix} y_1\cdots y_p\,\infty \\ y_{p+1}\cdots y_{2p+1}\end{matrix}\right\}.$$

We recall that the $d_{12\dots p}$ in (30) have been formed for $t_1,\cdots,t_{2p+1}$ and $t_{2p+2}=t$. Their polar coefficients have been formed from $y_1,\cdots,y_{2p+1},\infty$ which are projective to $t_1,\cdots,t_{2p+1},t$. Hence the canonizant is the form taken by the incidence condition (19) [cf. also (18)]. This leads to a determination of the nature of the p-ic spread (30), and of the place of the linear space ξ in the mapping of $S_{2p-1}(y)$ on $M_{2p-1}(x)$ in $S_{\nu-1}$. When $t_1,\cdots,t_{2p+1}$ are given, the norm-curve N^{2p-1} in $S_{2p-1}(y)$ is determined, and for given y in (27) the parameters t of the $2p-1$ spaces of N^{2p-1} on y are determined. The canonizant of this $(2p-1)$-ic furnishes the p values $t=\tau_1,\cdots,\tau_p$ such that y is on an S_{p-1} which cuts N^{2p-1} at the points $t=\tau_1,\cdots,\tau_p$, say *a p-secant* S_{p-1} of N^{2p-1}. If then this canonizant is set equal to zero, for given t we obtain the locus in variables y of the ∞^{p-1} p-secant S_{p-1}'s of N^{2p-1} which are on the point t of N^{2p-1}. Hence

(31) *In* (18) *the ordered* $(2p+2)$*-ic determines an incident point* x *and space* ξ *in* $S_{\nu-1}$. *When the* $(2p+2)$*-ic is first mapped on a point* y *in* $S_{2p-1}(y)$, *and* y *is then mapped upon* x *in* $S_{\nu-1}$ [cf. § 2 (7)], *the linear space* ξ *arises from a particular member of the mapping system defined at* y *by the basic norm-curve* N^{2p-1} *on* y. *This member is a cone of order* p, *the locus of the* ∞^{p-1} p*-secant spaces* S_{p-1} *of* N^{2p-1}, *each of which is on* y. *A section of this cone is the locus in* S_{2p-2} *of the* ∞^{p-1} $(p-1)$*-secant spaces* S_{p-2} *of a norm-curve* N^{2p-2}.

6. Applications to W_3 and W_p. We close this initial article with a contrast of the incidence condition described in § 5 (18), (31) for the two cases $p=2$ and $p=3$. This condition is bilinear in the invariants (A) and (B). The invariants (B) however can be expressed as polynomials of degree p in the invariants (A) [cf. [6]]. When the invariants (A) are expressed as polynomials of degree p in the coördinates of a point y in $S_{2p-1}(y)$, and the invariants (B) in terms of a point y' in $S_{2p-1}(y)$, this incidence condition becomes

$$f(y^p, y'^{p^2}) = 0, \tag{32}$$

where f is a double form of the degrees indicated. According to (31) this equation expresses that, for given y', y is a point on a p-secant space S_{p-1} through y' of the norm-curve N^{2p-1} determined by the set of $2p+2$ points P'^{2p-1}_{2p+2}, made up of the base of the mapping system (say $p_1, \cdots, p_{2p+1}$) and y'. In the notation introduced in connection with the generalized Weddle W_p ([4], p. 449), and for $y' = p_{2p+2}$, $f=0$ is the equation of the P-locus, $\pi_{1,2,\ldots,2p+1}$.

For $p=2$, the incidence condition, $f(y^2, y'^4)=0$, is for given y' a quadric cone with generators on y' and the points of the N^3 on $P'^3_6 = p_1, \cdots, p_5, y'$. Thus it is the condition that $P^3_6 = p_1, \cdots, p_5, y$, be on a quadric cone with vertex at y'. Hence for given y it is the equation in variables y' of the Weddle surface determined by P^3_6. The symmetry of the condition in the six points of P^3_6 is notable.

For $p=3$, this symmetry does not persist. The $f(y^3, y'^9)=0$ is, for given y', the locus of trisecant planes on y' of the N^5 on $P'^5_8 = p_1, \cdots, p_7, y'$. Therefore it is the condition that $P^5_8 = p_1, \cdots, p_7, y$ be projected from y' into $Q^4_8 = q_1, \cdots, q_7, z$ where z is a point on the bisecant locus B of the N^4 determined by $q_1, \cdots, q_7$. We have obtained [[4], p. 478 (67e)] the equation of B in terms of $q_1, \cdots, q_7$ in the form,

$$\begin{aligned}(33)\qquad &\Sigma(14567)(23567)(24157)(24367)\\ &\times(13257)(13467)(2413z)(2456z)(1356z)=0,\end{aligned}$$

where Σ refers to the six terms obtained by permutation of 1, 2, 6 (or 3, 4, 5). By the Clebsch principle of transference,

$$(34) \qquad f(y^3, y'^9) = \Sigma(14567y')(23567y')(24157y')(24367y') \times (13257y')(13467y')(2413yy')(2456yy')(1356yy') = 0.$$

Thus the condition $f(y^3, y'^9) = 0$ is of degree six in $p_1, \cdots, p_7$ and of degree three in y.

For $p = 2$ and $P_6{}^3 = p_1, \cdots, p_5, p_6 = y$, the surface $f(y^2, y'^4) = 0$ is precisely the Weddle W_2. For $p = 3$ and $P_8{}^5 = p_1, \cdots, p_7, p_8 = y$, we ask whether the spread $f(y^3, y'^9) = 0$ contains the Weddle 3-way, W_3; and, if not, how it cuts W_3. If we follow out the argument ([4], pp. 478-9) from which the equation (33) was obtained, in the form in which it transfers, as in (34), to S_5, we obtain

$$(35) \qquad H_{24}H_{13}G_{148}G_{238} - H_{23}H_{14}G_{138}G_{248} = G_{567} \cdot f(y^3, y'^9).$$

In this G_{123} is the S_4 on $p_4, \cdots, p_8$, and H_{78} is the 4-way Weddle cone of order 4 with 4-fold line p_7p_8 obtained by the Clebsch principle from the W_2 on points $q_1, \cdots, q_6$ in S_3. With respect to the indices $1, \cdots, 6$; 7, 8 the F-loci of the third kind of $P_8{}^5$ [cf. [4], § 1] divide into types $\pi^{(3)}_{178}$, $\pi^{(3)}_{128}$, $\pi_8{}^{(3)}$, $\pi_1{}^{(3)}$ where $\pi^{(3)}_{ijk}$ is the plane $p_ip_jp_k$, and $\pi_i{}^{(3)}$ is the quartic 2-way cone of bisecants of N^5 on p_i. Then H_{78} contains $\pi^{(3)}_{178}$ doubly, $\pi^{(3)}_{128}$ and $\pi_8{}^{(3)}$ simply, and does not contain $\pi_1{}^{(3)}$ [cf. [4] (78f)]. Also H_{78} contains the surface $V_2{}^{(5)}$ [cf. [4] (76)] with respect to which we wish to correct and amplify the theorem [[4] (82)] to read as follows:

(36) *W_3 contains the surface $V_2{}^{(5)}$, a 2-way of order 26 with 4-fold points at $P_8{}^5$ and with simple curve N^5 and simple lines p_ip_j.*

The correction is with respect to the order 26 which was given as 15. In fact $V_2{}^{(5)}$ is projected from p_8 into $V_2{}^{(4)}$ of order 11 doubly covered as will appear later in connection with (39). This double covering of $V_2{}^{(4)}$ was overlooked in ([4]).

In order to find the intersection of (35) with W_3 we examine first the intersection of H_{78} with W_3. Since W_3 has the order 19 with 9-fold points at $P_8{}^5$, triple curves N^5 and p_ip_j, and simple F-loci of the third kind of $P_8{}^5$ [[4] (81)], H_{78} cuts W_3 in a 2-way of order 76 from which the six $\pi^{(3)}_{178}$ separate doubly, the thirty $\pi^{(3)}_{128}$ and the two cones $\pi_7{}^{(3)}$, $\pi_8{}^{(3)}$ separate simply, the residual intersection is precisely $V_2{}^{(5)}$ of order 26 with 4-fold points at $P_8{}^5$ and simple curve N^5. Since $V_2{}^{(5)}$ is itself invariant under the Cremona group $G_{2.2^6}$ of W_3, it must also contain the lines p_ip_j conjugate to N^5 simply.

It is now clear from the form of the left member of (35) that $f(y^3, y'^9) = f(p_8^3, y'^9) = 0$ contains $V_2^{(5)}$ doubly, the 35 planes $\pi_{123}^{(3)}$ doubly, and the 21 planes $\pi_{128}^{(3)}$ as well as the seven cones $\pi_1^{(3)}$ simply. Since $2.26 + 35.2 + 21 + 7.4 = 9.19$, the total intersection of $f(p_8^3, y'^9) = 0$ with W_3 is accounted for. Moreover the left member of (35) does not contain W_3. For, according to [4 (78h)], the following octavic spreads contain W_3:

$$(37) \quad \pm H_{23}H_{14}/(23)(14) = \pm H_{31}H_{24}/(31)(24) = \pm H_{12}H_{34}/(12)(34).$$

Hence on W_3 this left member is equivalent to

$$(31)(24)G_{148}G_{238} \pm (23)(14)G_{138}G_{248} = G_{12348},$$

the quadric with nodes at p_5, p_6, p_7 and simple points at $p_1, \cdots, p_4, p_8$ [cf. 4 (78e)]. This quadric however does not contain W_3 but rather meets W_3 in certain F-loci of the third kind and in a 9-ic 2-way which is on G_{567} also [4, p. 489]. Hence

(38) *The 9-ic locus, $f(p_8^3, y'^9) = 0$, which is determined by the incidence condition of* § 5 (18), (31) *has a nodal locus $V_2^{(5)}$. It also contains all the F-loci of P_8^5 of the third kind except the quartic cone $\pi_8^{(3)}$, and in particular contains the* 35 *planes of type $\pi_{123}^{(3)}$ doubly.*

Thus the property of W_2 with respect to this incidence condition extends not to W_3 but rather to the surface $V_2^{(5)}$ on W_3. This is a situation which will frequently recur.

We have attached the ∞^p points of W_p to the ∞^p projectively distinct but birationally equivalent curves H_p^{p+2} of genus p and order $p+2$ with p-fold point at O and branch points $r_1, \cdots, r_{2p+2}$ for which the pencil of lines from O to the branch points is projectively given. We inquire therefore as to the projective peculiarity of H_p^{p+2} when the point of W_p is on $V_2^{(2p-1)}$.

In S_{2p-1} a normal elliptic curve E^{2p} depends upon $4p^2$ constants. Since it is $2p-2$ conditions that E^{2p} pass through a given point, there are ∞^4 curves E^{2p} on P_{2p+2}^{2p-1}. If such a curve E^{2p} has two nodes it breaks up into two rational norm-curves bisecant to each other. These degenerate E^{2p}'s are distributed in a variety of families, each ∞^2. A particular set of $2p+2$ of the families is of special interest. One such family consists of ∞^2 curves each made up of a line on p_i and an N^{2p-1} bisecant to the line and on the remaining points of P_{2p+2}^{2p-1}. As i runs from 1 to $2p+2$ the $2p+2$ families are obtained. Each of these $2p+2$ families has the property that it is invariant under the Cremona group $G_{2.2^{2p}}$ of W_p. This indeed is almost evident for the generators I_{ij} of $G_{2.2^{2p}}$. If n_1, n_2 are the nodes of an E^{2p} in

the family determined by p_i, then from n_1 (or n_2) the set P_{2p+2}^{2p-1} is projected into a set Q_{2p+2}^{2p-1} in S_{2p-2} which lies on an N^{2p-2}. But this is the condition that n_1 be a point on $V_2^{(2p-1)}$ [cf. [4] (76)].

The existence of this family of degenerate E^{2p}'s explains the double covering of the projection of $V_2^{(2p-1)}$ upon $V_2^{(2p-2)}$ from the point p_i. By beginning with a point of $V_2^{(2p-1)}$ and projecting it, and the new points so obtained, from the points of P_{2p+2}^{2p-1} upon $V_2^{(2p-1)}$ a closed system of 2.2^{2p} points is obtained on $V_2^{(2p-1)}$ analogous to Baker's closed set of 2.16 points on W_2. The closure is a consequence of the invariance of each of the $2p+2$ families under the group of W_p.

(39) *A line through the $2(p-1)$-fold point p_i of $V_2^{(2p-1)}$ to a point of $V_2^{(2p-1)}$ meets $V_2^{(2p-1)}$ in another point. By repeated projections from points p_i a point y on $V_2^{(2p-1)}$ gives rise to a closed set of 2.2^{2p} points on $V_2^{(2p-1)}$ which divide into two sets of 2^{2p} points, each a conjugate set under the group of the W_p.*

Here we again have a property of W_2 which generalizes into a property of $V_2^{(2p-1)}$ rather than a property of W_p.

The essential condition on a point y of $V_2^{(2p-1)}$ is that the set P_{2p+2}^{2p-1} is projected from y upon a set Q_{2p+2}^{2p-2} on an N^{2p-2}. But P_{2p+2}^{2p-1}, y are associated with the planar set of points $r_1, \cdots, r_{2p+2}, O$. Hence the projection, Q_{2p+2}^{2p-2}, of P_{2p+2}^{2p-1} from y is associated with the planar set $r_1, \cdots, r_{2p+2}$ [1, I, p. 158 (6)]. But Q_{2p+2}^{2p-2} is a set on a rational norm-curve N^{2p-2} whence $r_1, \cdots, r_{2p+2}$ is also a set on a rational norm-curve, i. e., a conic, with parameters on the conic projective to the parameters of Q_{2p+2}^{2p-2} on N^{2p-2} [7, p. 12, Theo. (10)]. If $H_p{}^{p+2} \equiv x_0{}^2 f_p + 2x_0 f_{p+1} + f_{p+2} = 0$, the branch points of $H_p{}^{p+2}$ are on the conic $H_0{}^2 \equiv x_0{}^2 g_0 + 2x_0 g_1 + g_2 = 0$ if

$$g_0 f_{p+2} - 2g_1 f_{p+1} + g_2 f_p \equiv 0. \tag{40}$$

This identical vanishing of the $(p+2)$-ic (40) imposes $p+3$ conditions, of which only five can be satisfied by proper choice of g_0, g_1, g_2, leaving $2p-2$ projective conditions to be satisfied by f_p, f_{p+1}, f_{p+2}. These conditions confine the representative point y on W_p to the locus $V_2^{(2p-1)}$. Hence

(41) *The locus $V_2^{(2p-1)}$ on W_p is that locus whose points correspond to hyperelliptic curves $H_p{}^{p+2}$ whose $2p+2$ branch points are on a conic.*

It may be observed that, for $p=3$, the 9 points $r_1, \cdots, r_8, O$ are the base points of a pencil of cubics. Thus O is determined in general by

$r_1, \cdots, r_8$. If however $r_1, \cdots, r_8$ are on a conic, O may be any point of the plane. Let $s_1, \cdots, s_8$ be the parameters of $r_1, \cdots, r_8$ on their conic. In S_5 let $s_1, \cdots, s_7, s_8$ be the parameters of the points $p_1, \cdots, p_7, n_2$ on the N'^5 through them. On the quartic 2-way cone of lines from n_2 to the points of N'^5 let p_8 be any point, and let the line $p_8 n_2$ cut N'^5 again at n_1. Then the Weddle W_3 determined by $p_1, \cdots, p_8$ will contain a $V_2^{(5)}$ on n_1, n_2. On the N^5 determined by $p_1, \cdots, p_8$ these points will have parameters $t_1, \cdots, t_8$. Then there will be a unique point O in the plane for which the octavic pencil of lines from O to $r_1, \cdots, r_8$ will be projective to $t_1, \cdots, t_8$. As p_8 runs over the above quartic cone, the point O runs over the plane.

On W_p there will be similar manifolds of higher dimension, V_4, $V_6, \cdots$, which map curves H_p^{p+2} whose branch points are respectively on a cubic with O, on a quartic with node at $O, \cdots$. Doubtless these further manifolds in the higher cases will have properties as interesting as those of $V_2^{(2p-1)}$.

REFERENCES

[1] A. B. Coble, "Point Sets and Allied Cremona Groups," *Transactions of the American Mathematical Society*, I, Vol. 16 (1915), pp. 155-198; II, Vol. 17 (1916), pp. 345-385; III, Vol. 18 (1917), pp. 331-372.

[2] A. B. Coble, "Finite Geometry . . . Characteristic Theory of the Odd and Even Theta Functions," *Transactions of the American Mathematical Society*, Vol. 14 (1913), pp. 241-276.

[3] A. B. Coble, "Algebraic Geometry and Theta Functions," *Colloquium Lectures of the American Mathematical Society*, Vol. 10, New York (1929).

[4] A. B. Coble, "A Generalization of the Weddle Surface . . .," *American Journal of Mathematics*, Vol. 52 (1930), pp. 439-500.

[5] W. Wirtinger, "Untersuchung über Thetafunctionen," Leipzig (Teubner, 1895).

[6] C. M. Huber, "On Complete Systems of Irrational Invariants of Associated Point Sets," *American Journal of Mathematics*, Vol. 49 (1927), pp. 251-267.

[7] A. B. Coble, "Associated Sets of Points," *Transactions of the American Mathematical Society*, Vol. 24 (1922), pp. 1-20.

ON THE TOPOLOGY OF ALGEBROID SINGULARITIES.

By OSCAR ZARISKI.

The classification of algebroid singularities from a topological point of view has been the object of a thorough investigation by K. Brauner.* If the singularity consists of one branch, given by a Puiseux expansion $y = y(x)$, its intersection with the boundary of a small neighborhood, say of the 4-cell $|x| < \text{const.}$, $|y| < \text{const.}$, is projected stereographically into a knot of the ordinary space. Brauner gives a full description of these knots and he also derives the generating relations of their fundamental group.†

If the singularity consists of several branches, the above intersection consists of several linked knots.

To complete the classification of the algebroid singularities from a topological point of view it would be still necessary to prove that the different knots thus obtained are actually distinct, i. e., are not isotopic. The case of singular branches of genus 1, which give rise to knots lying on a torus, is settled by a paper by O. Schreier.‡ The purpose of the present paper is to prove the general theorem: §

* K. Brauner, "Zur Geometrie der Funktionen zweier Veränderlichen": III. Klassifikation der Singularitäten algebroider Kurven; IV. Die Verzweigungsgruppen; *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 6 (1928), pp. 8-54.

† It is apparent, however, that this part of Brauner's paper, though extremely interesting in its conclusions, is unnecessarily long and complicated. E. Kähler obtains Brauner's results in a simpler manner in his paper "Über die Verzweigung einer algebraischen Funktion zweir Veränderlichen in der Umgebung einer singulären Stelle," *Mathematische Zeitschrift*, Vol. 30 (1929), pp. 188-204.

‡ O. Schreier, "Über die Gruppen $A^aB^b = 1$," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 3 (1924), pp. 167-169.

§ After having communicated this theorem at a meeting of the American Mathematical Society in March, 1932, I learnt from Professor Lefschetz that the same theorem has been proved by Werner Burau, and that Burau's paper entitled "Kennzeichnung der Schlauchknoten" is being published in the forthcoming issue of the *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität.* The polynomial $F(x)$ in section (5) of this paper is also obtained by Burau as the Alexander polynomial of the knot (See J. W. Alexander, "Topological Invariants of Knots and Links," *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 273-306). This was to be expected since the connection between the homology characters of the k-sheeted manifolds with the knot as branch curve and the Alexander polynomial of the knot was clearly pointed out in general by Alexander in his quoted paper.

Conceding priority to Mr. Burau, I publish only an outline of my proof. I do

If two algebroid singularities, each composed of a single branch, are distinct from an algebro-geometric point of view, i. e., are either of distinct genera, or being of the same genus do not have the same pairs of characteristic numbers, they are also topologically distinct, and their fundamental groups are not isomorphic.*

The method of proof is based on the consideration of k-sheeted *cyclic Riemann manifolds* having the given singularity as locus of branch points. The idea of using these manifolds as a source of invariants has been applied by Reidemeister † in his treatment of the theory of knots and also by the author in two papers dealing with the fundamental group of plane algebraic curves.‡

It seemed advisable to the author to include in this paper another method of deriving the generating relations of the fundamental group of a singularity composed of a single branch. This method differs from the one used by Brauner in that it is algebraic and operates directly on the given singularity instead of on the corresponding knot. It can be readily extended to singularities composed of several branches.

The Fundamental Group.

1. *Preliminary remarks.* Any closed path in the residual space of the given singularity $y = y(x)$, can be deformed, one of its points remaining fixed, into a circuit lying in a "line", $x = \text{const.}$§ If the order of the branch

---

this especially because of the additional result arrived at in the course of the proof to the effect that the Betti number R_1 of the k-sheeted Riemann manifold is the number of roots of the polynomial $F(x)$ which are also roots of the equation $x^k = 1$. Whether this is a general property of the Alexander polynomials or holds only for the particular knots connected with algebraic singularities—is a question which would be worthwhile considering.

* The characteristic numbers of a singular branch determine completely the character of the singularity in the sense of Nöther, i. e., as an aggregate of successive, infinitely near multiple points. For this and related notions of the theory of singularities see the original exposition in F. Enriques and O. Chisini, *Teoria geometrica delle equazioni e delle funzioni algebriche*, Vol. 2, book 4, Ch. I.

† "Knoten und Gruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, Vol. 6 (1926), pp. 7-23.

‡ "On the Linear Connection Index of the Algebraic Surfaces $z^n = f(x, y)$," *Proceedings of the National Academy of Sciences*, Vol. 15, No. 6 (1929), pp. 494-501; "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, series 2, Vol. 32, No. 3 (1931), pp. 485-511.

§ See S. Lefschetz, *L'Analysis Situs et la géometrie algébrique*, p. 33. See also my paper "On the Problem of Existence of Algebraic functions of Two Variables Possessing a Given Branch Curve," *American Journal of Mathematics*, Vol. 51, No. 2 (April, 1929), where the method outlined below has been applied to the similar problem at large, involving algebraic curves instead of the neighborhood of an algebroid singularity.

is n, that circuit is a product of loops $g_0, g_1, \cdots, g_{n-1}$, each surrounding one of the points $y_0, y_1, \cdots, y_{n-1}$, in which that line intersects the branch. Hence the loops g_i can be taken as generators of the fundamental group of the given singularity. As x makes one turn around the origin $x = 0$, (the fixed initial point of the considered circuits being the point at infinity on the y-axis), the original set of non-intersecting loops g_i is continuously deformed into a new set of non-intersecting loops g'_i. We obtain all the generating relations of the fundamental group by expressing the loops g'_i in terms of the loops g_i, and by putting $g'_i = g_i (i = 0, 1, 2, \cdots, n-1)$.*

Let
$$y = \sum_{\nu=1}^{\infty} a_\nu x^{\alpha_\nu/n},$$

be the Puiseux expansion of the given branch, which we shall denote by Γ. We suppose that the axis $y = 0$ is tangent to the branch. Then $\alpha_1 > n$. Let $\alpha_1/n = m_1/n_1$, where m_1 and n_1 are relatively prime integers, in symbols $(m_1, n_1) = 1$. We write the exponent α_2/n of the second term of the series in the following form:

$$\alpha_2/n = m_1/n_1 + m_2/n_1 n_2,$$

where $(m_2, n_2) = 1$; and more in general, let

$$\alpha_\nu/n = m_1/n_1 + m_2/n_1 n_2 + \cdots + m_\nu/n_1 n_2 \cdots n_\nu,$$

where $(m_\nu, n_\nu) = 1$. Obviously there is only a finite number of values of ν, for which $n_\nu > 1$. If n_q is the last n_ν, which is greater than 1, then $n_1 n_2 \cdots n_q = n$. The terms $\alpha_\nu x^{\alpha_\nu/n}$, for which $n_\nu > 1$, are called the *characteristic terms* of the Puiseux expansion, and the number p of such terms is called the *genus* of the branch.

It is well known that the type of the given singularity, from the standpoint of analytic transformations on the variables x and y, is completely determined by the characteristic terms defined above. It follows *a fortiriori* (and this has also been proved directly by Brauner) that also topologically the given algebroid branch is completely determined by the exponents of the characteristic terms. We therefore replace the given series by the sum of its characteristic terms:

---

* The assertion that we obtain in this manner *all* the generating relations of the group follows from the fact that the relations $g'_i = g_i$ offer the sufficient conditions for the existence of functions of two variables possessing the given singularity as locus of branch points. For the proof of this last statement see F. Enriques, "Sulla costruzione delle funzioni algebriche di due variabili possedenti una data curva di diramazione," *Annali di Matematica pura ed applicata*, Ser. 4, Vol. 1 (November, 1923), pp. 185-198.

$$y = a_1 x^{m_1/n_1} + a_2 x^{m_1/n_1+m_2/n_1n_2} + \cdots + a_p x^{m_1/n_1+m_2/n_1n_2+\cdots+m_p/n_1n_2\cdots n_p},$$

where $n_i > 1$, $(i = 1, 2, \cdots, p)$, and $(m_i, n_i) = 1$. Since the actual values of the coefficients are irrelevant, we shall put for conveniency $a_i = 1/2^{i-1}$.

We shall derive the fundamental group of the given branch Γ by a procedure of induction, considering first the case of branches of genus 1 and 2.

2. *The fundamental group of the branch* $y = x^{m_1/n_1}$. We let x vary on the circle $|x| = 1$ starting from the initial value $x = 1$. We mark in the plane of the complex variable y the initial values of y:

$$y_0 = 1,\ y_1 = e^{2\pi i/n_1}, \cdots, y_{n_1-1} = e^{2\pi i(n_1-1)/n_1},$$

and we fix the loops g_i so that if $m_1 = 1$, then as x makes one turn about the origin, the loops g_i should be transformed as follows:

$$(1) \qquad g'_0 = g_1,\ g'_1 = g_2, \cdots, g'_{n_1-2} = g_{n_1-1},\ g'_{n_1-1} = P_1^{-1} g_0 P_1,$$

where $P_1 = g_0 g_1 \cdots g_{n_1-1}$ is the circuit surrounding all the points y_i. We shall denote this transformation by T_1. It is obvious that P_1 is an invariant circuit: $T_1(P_1) = P_1$.

For an arbitrary m_1 the transformation of the loops obtained by letting x turn around the origin coincides with the m_1-th power of T_1. If we put

$$m_1 = h_1 n_1 + l_1,$$

where $n_1 > l_1 \geqq 0$, then $T_1^{m_1}$ can be written as follows:

$$g'_0 = P_1^{-h_1} g_{l_1} P_1^{h_1},\ g'_1 = P_1^{-h_1} g_{l_1+1} P_1^{h_1}, \cdots, g'_{n_1-l_1-1} = P_1^{-h_1} g_{n_1-1} P_1^{h_1};$$
$$g'_{n_1-l_1} = P_1^{-h_1-1} g_0 P_1^{h_1+1}, \cdots, g'_{n_1-1} = P_1^{-h_1-1} g_{l_1-1} P_1^{h_1+1}.$$

We therefore obtain for the fundamental group, which we shall denote by G_1, the following generating relations:

$$(2) \qquad \begin{aligned} g_{i+l_1} &= P_1^{h_1} g_i P_1^{-h_1}, && (i = 0, 1, \cdots, n_1 - l_1 - 1), \\ g_i &= P_1^{h_1+1} g_{n_1-l_1+i} P_1^{-h_1-1}, && (i = 0, 1, \cdots, l_1 - 1). \end{aligned}$$

Hence G_1 is generated by the two elements g_0 and P_1, since $n_1 - 1$ of the above relations can be used for expressing the remaining g_i's in terms of g_0 and P_1. By elimination we derive the following relation between g_0 and P_1:

$$(3) \qquad g_0 = P_1^{m_1} g_0 P_1^{-m_1},$$

which could have also been obtained by operating directly on g_0 by the transformation $T_1^{m_1 n_1}$. We obtain a second generating relation by replacing in $P_1 = g_0 g_1 \cdots g_{n_1-1}$ the factors $g_1, g_2, \cdots, g_{n_1-1}$ by their expressions in terms

of g_0 and P_1, derived from (2). We denote by x_1 and y_1 two positive integers such that

$$x_1 m_1 = y_1 n_1 + 1, \tag{3'}$$

and we consider the relations between the g_i's to which the transformation $T_1^{m_1 x_1}$ gives rise. These relations, which are obviously a consequence of (2), are the following:

$$\begin{aligned} g_{i+1} &= P_1^{y_1} g_i P_1^{-y_1}, && (i = 0, 1, \cdots, n-2), \\ g_0 &= P_1^{y_1+1} g_{n_1-1} P_1^{-y_1-1}. \end{aligned} \tag{4}$$

Conversely, the relations (2) follow from (3) and (4). In fact, from (4) we derive

$$\begin{aligned} g_{i+l_1} &= P_1^{y_1 l_1} g_i P_1^{-y_1 l_1}, && (i = 0, 1, \cdots, n_1 - l_1 - 1), \\ g_i &= P_1^{y_1 l_1 + 1} g_{n_1 - l_1 + i} P_1^{-y_1 l_1 - 1}, && (i = 0, 1, \cdots, l_1 - 1). \end{aligned}$$

Now these relations lead back to (2), since by (3) and (4) $P_1^{m_1}$ is commutative with each element g_i, and since moreover

$$(x_1 l_1 - 1) m_1 = (y_1 l_1 - h_1) n_1, \text{ i. e., } y_1 l_1 \equiv h_1 \pmod{m_1}.$$

From (4) we deduce $g_{i+1} = P_1^{i y_1 l_1} g_0 P_1^{-i y_1 l_1}$ $(i = 0, 1, \cdots, n_1 - 2)$ and substituting into $P_1 = g_0 g_1 \cdots g_{n_1-1}$, we find the required second relation between g_0 and P_1:

$$(P_1^{y_1} g_0)^{n_1} = P_1^{m_1 x_1}. \tag{5}$$

Let $$Q_1 = (P_1^{y_1} g_0)^{m_1} P_1^{-m_1 y_1}.$$

Then, by (3) and (5),

$$Q_1^{x_1} = (P_1^{y_1} g_0)^{m_1 x_1} P_1^{-m_1 x_1 y_1} = (P_1^{y_1} g_0)(P_1^{y_1} g_0)^{n_1 y_1} P_1^{-m_1 x_1 y_1},$$

or

$$Q_1^{x_1} = P_1^{y_1} g_0, \tag{6}$$

and hence P_1 and Q_1 can be taken as generators of G_1. The relations (3) and (5) lead to the following relation between P_1 and Q_1:

$$Q_1^{n_1} = P_1^{m_1}, \tag{7}$$

and conversely from this relation follow (3) and (5), if we use (6) to define g_0 in terms of P_1 and Q_1. Consequently, *the group G_1 is generated by two elements P_1 and Q_1, satisfying the relation* (7).

3. *The fundamental group G_2 of a branch of genus 2:*

$$y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1 + m_2/n_1 n_2}. \tag{8}$$

Let, as before, $y_k = e^{2\pi ik/n_1}$, $(k = 0, 1, \cdots, n_1 - 1)$. Then the $n_1 n_2$ values of y in (8), for $x = 1$, can be denoted by y_{ij}, where i runs from 0 to $n_1 - 1$ and j runs from 0 to $n_2 - 1$, in such a manner that for a fixed i the n_2 points y_{ij} lie on the circle of center y_i and of radius 1/2. We choose our notations so that arg. $(y_{i,j+1} - y_i)/(y_{ij} - y_i) = 2\pi i/n_2$, and that $y_{00} = 3/2$. It is immaterial, for the sequel, which of the points y_{ij}, for i fixed and > 0, is denoted by y_{i0}.

We shall accordingly denote the loop surrounding the point y_{ij} by g_{ij}. We choose these loops in the following manner: We first construct the loops g_i $(i = 0, 1, 2, \cdots, n_1 - 1)$, considered in the preceding case, where now g_i surrounds the points $y_{i0}, y_{i1}, \cdots, y_{i,n_2-1}$, and hence also y_i. We then choose the loops g_{ij} in such a manner that

$$g_{i0} g_{i1} \cdots g_{i,n_2-1} = g_i.$$

Let T_2 be the transformation of the loops g_{ij} affected by a complete turn of the point x around the origin. This transformation induces a transformation of the loops g_i, which coincides with the transformation T_1 considered in the preceding case. We therefore have, as a first set of generating relations of G_2, the generating relations (2) of G_1, or the relation (7) involving the reduced set of generators P_1 and Q_1. As for the new relations between the g_{ij}'s, it is not necessary to write them all explicitly, as most of them serve only to express the g_{ij}'s in terms of the g_i's and of the g_{0j}'s. For instance, the loops $g_{0,j}$ go by T_2 into transformed of the loops $g_{l_1,j}$, where l_1 is defined by (3′), and the corresponding generating relations are of the type:

$$g_{l_1,j} = A_j g_{0,j'} A_j^{-1},$$

where the elements A_j are in G_1, i. e., are expressible in terms of the g_i's, hence also in terms of P_1 and Q_1. If we express the fact that $\prod_{j=0}^{n_2-1} A_j g_{0,j'} A_j^{-1} = g_{l_1}$, we obtain a relation in which the loops $g_{0,j'}$ will necessarily combine into their product g_0, and the result will be obviously one of the relations (2), namely: $g_{l_1} = P_1^{h_1} g_0 P_1^{-h_1}$.

In a similar manner we express all the loops g_{ij} in terms of the loops $g_{0,j}$, g_i, by considering the successive powers $T_2^2, T_2^3, \cdots, T_2^{n_1-1}$ of T_2.

In order to obtain relations for this reduced set of generators of G_2, we must consider the transformation $T_2^{n_1}$, obtained by letting the variable x make n_1 complete turns around the origin. We put $x = z^{n_1}$, and hence

$$y = z^{m_1} + 1/2\, z^{m_1 + m_2/n_2},$$

and we let z make one turn around $z = 0$. If we put

$$m_2 = h_2 n_2 + l_2, \tag{9}$$

we easily conclude that the transformation $T_2^{n_1}$ has the following form:

$$T_2^{n_1} \equiv \begin{cases} g'_{0j} = g_0^{-h_2} P_1^{-m_1} g_{0,j+l_2} P_1^{m_1} g_0^{h_2}, & (j = 0, 1, \cdots, n_2 - l_2 - 1); \\ g'_{0,n_2-l_2+j} = g_0^{-h_2-1} P_1^{-m_1} g_{0,j} P_1^{m_1} g_0^{h_2+1}, & (j = 0, 1, \cdots, l_2 - 1). \end{cases}$$

We have then, in addition to the relation (7), the following relations:

$$\begin{cases} g_{0,j+l_2} = P_1^{m_1} g_0^{h_2} g_{0,j} g_0^{-h_2} P_1^{-m_1} & (j = 0, 1, \cdots, n_2 - l_2 - 1); \\ g_{0,j} = P_1^{m_1} g_0^{h_2+1} g_{0,n_2-l_2+j} g_0^{-h_2-1} P_1^{-m_1}, & (j = 0, 1, \cdots, l_2 - 1). \end{cases} \tag{10}$$

The relations (7) and (10) constitute a complete set of generating relations of the group G_2. Obviously a further reduction of the number of generators is possible, since the elements $g_{0,j}$, $j > 0$, can be eliminated, leaving g_{00}, P_1 and Q_1 as only generators of G_2. In order to accomplish this elimination we proceed in a manner similar to the one followed in the preceding case. Observing that by (3) g_0 is commutative with $P_1^{m_1}$, we derive from (10) the following relation:

$$g_{00} = g_0^{m_2} P_1^{m_1 n_2} g_{00} (g_0^{m_2} P_1^{m_1 n_2})^{-1}, \tag{11}$$

which could have also been derived by operating directly on g_{00} by $T_2^{n_1 n_2}$. Let x_2 and y_2 be positive integers, such that

$$x_2 m_2 = y_2 n_2 + 1. \tag{12}$$

The transformation $T_2^{n_1 x_2}$ leads to relations which are obviously a consequence of the relations (10). These relations are:

$$\begin{aligned} &g_{0,j+1} = B g_{0,j} B^{-1}, && (j = 0, 1, \cdots, n_2 - 2), \\ &g_{0,0} = B g_0 g_{0,n_2-1} g_0^{-1} B^{-1}, \end{aligned} \tag{13}$$

where $B = g_0^{y_2} P_1^{m_1 x_2}$. Conversely, the relations (10) follow from these relations and from the relation (11). In fact, from (13) we deduce

$$\begin{aligned} &g_{0,j+l_2} = B^{l_2} g_{0,j} B^{-l_2}, && (j = 0, 1, \cdots, n_2 - l_2 - 1), \\ &g_{0,j} = B^{l_2} g_0 g_{0,n_2-l_2-1} g_0^{-1} B^{-l_2}, && (j = 0, 1, \cdots, l_2 - 1). \end{aligned} \tag{14}$$

Now, we have the following identity: $(y_2 l_2 - h_2) n_2 = (x_2 l_2 - 1) m_2$, and therefore $y_2 l_2 - h_2 = m_2 k$, $x_2 l_2 - 1 = n_2 k$, where k is an integer. Consequently, using throughout the fact that g_0 and $P_1^{m_1}$ are commutative, we find:

$$B^{l_2} = (g_0^{m_2} P_1^{m_1 n_2})^k g_0^{h_2} P_1^{m_1}.$$

Since, by (11), $g_0^{m_2} P_1^{m_1 n_2}$ is commutative with g_{00} and therefore also, by (13), with any of the elements $g_{0,j}$, we see that the relations (10) are a consequence of the relations (14) and (11).

Using (13) in order to express all the elements $g_{0,j}$ in terms of P_1, g_0 and g_{00}, and substituting into the relation $g_0 = g_{00} g_{01} \cdots g_{0,n_2-1}$, we find:

$$(11') \qquad (g_{00}g_0^{y_2}P_1^{m_1x_2})^{n_2} = g_0^{m_2x_2}P_1^{m_1n_2x_2}.$$

The relations (11) and (11′) are, together with (7), the generating relations of group G_2. It is possible to replace the relations (11) and (11′) by one relation. In fact, let us consider the following element of G_2:

$$Q_2 = (g_{00}g_0^{y_2}P_1^{m_1x_2})^{m_2}(g_0^{m_2}P_1^{m_1n_2})^{-y_2}.$$

It is then easily seen that

$$(15) \qquad Q_2^{x_2} = g_{00}g_0^{y_2}P_1^{m_1x_2},$$

and consequently we can use Q_2 as one of the generators of G_2, instead of g_{00}. The relations (11) and (11′) lead to the following:

$$Q_2^{n_2} = g_0^{m_2}P_1^{m_1n_2} = g_0^{m_2}Q_1^{n_1n_2},$$

and conversely (11) and (11′) follow from this relation, if we use (15) in order to define g_{00} in terms of P_1, g_0 and Q_1. We change slightly our notations, using the letters P_2 and P_3 in order to denote the elements g_0 and g_{00} respectively. With these notations we may state the final result as follows: *The group G_2 is generated by three elements P_1, Q_1 and Q_2, satisfying the following two relations:*

$$Q_1^{n_1} = P_1^{m_1}, \quad Q_2^{n_2} = P_2^{m_2}Q_1^{n_1n_2},$$

where $P_2 = Q_1^{x_1}P_1^{-y_1}$. The elements P_1, P_2 and P_3 have the following significance: P_3 is a loop surrounding one point, y_{00}, P_2 surrounds n_2 points y_{0j}, and P_1 surrounds all the n_1n_2 points y_{ij}.

4. *The fundamental group G_p of a branch Γ_p of genus p:*

$$(16) \qquad y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1+m_2/n_1n_2} + \cdots + 1/2^{p-1}x^{m_1/n_1+m_2/n_1n_2+\ldots+m_p/n_1n_2\cdots n_p}.$$

We prove that: *The group G_p is generated by $p+1$ elements: P_1, Q_1, $Q_2, \cdots, Q_p$. The generating relations of G_p are the following:*

$$(17) \qquad Q_i^{n_i} = P_i^{m_i}Q_{i-1}^{n_{i-1}n_i}, \qquad (i = 1, 2, \cdots, p;\ Q_0 = 1),$$

where the elements $P_2, P_3, \cdots, P_p$ are defined by the following relations: *

$$(17') \qquad P_{i+1}P_i^{y_i}Q_{i-1}^{n_{i-1}x_i} = Q_i^{x_i}, \qquad (i = 1, 2, \cdots, p-1).$$

Here x_i and y_i are positive integers, such that

$$x_im_i = y_in_i + 1.$$

Moreover, if an element P_{p+1} is defined as follows:

$$(17'') \qquad P_{p+1}P_p^{y_p}Q_{p-1}^{n_{p-1}x_p} = Q_p^{x_p},$$

* See K. Brauner, *loc. cit.*

the elements $P_1, P_2, \cdots, P_{p+1}$ have the following significance: P_i is a loop surrounding $n_i n_{i+1} \cdots n_p$ of the values of the function y for $x = 1$ (one value, if $i = p+1$).

Since the theorem has already been proved for $p = 1, 2$, we prove it by induction, assuming that it is true for $p-1$, and namely for the branch Γ_{p-1}:

$$y = x^{m_1/n_1} + 1/2\, x^{m_1/n_1 + m_2/n_1 n_2} + \cdots + 1/2^{p-2} x^{m_1/n_1 + m_2/n_1 n_2 + \cdots + m_{p-1}/n_1 n_2 \cdots n_{p-1}},$$

obtained by dropping the last term in (16). We have then for the group G_{p-1} of this branch the generators $P_1, Q_1, Q_2, \cdots, Q_{p-1}$, and the generating relations:

$$Q_i^{n_i} = P_i^{m_i} Q_{i-1}^{n_{i-1} n_i}, \qquad (i = 1, \cdots, p-1), \tag{18}$$

where

$$P_{i+1} P_i^{y_i} Q_{i-1}^{n_{i-1} x_i} = Q_i^{x_i}, \qquad (i = 1, 2, \cdots, p-1). \tag{18'}$$

The elements P_i are loops in the plane, $x = 1$, of the variable y and surround $n_i n_{i+1} \cdots n_{p-1}$ points y_i of the branch Γ_{p-1}. On each circle of center y_i and radius $1/2^{p-1}$ there are n_p points y_{ij} of the branch Γ_p. The elements P_i can therefore be considered as elements of G_p, each P_i surrounding $n_i n_{i+1} \cdots n_p$ points y_{ij} of Γ_p. Let the notations be so chosen that $y_0 = 1 + 1/2 + \cdots + 1/2^{p-2}$ and $y_{00} = y_0 + 1/2^{p-1}$, and let $g_0, g_1, \cdots, g_{n_p - 1}$ be a set of loops surrounding the points $y_{00}, y_{01}, \cdots, y_{0, n_p - 1}$ respectively, and such that

$$g_0 g_1 \cdots g_{n_p - 1} = P_p.$$

Then it follows as in the preceding case $p = 2$, that the elements g_j together with the generators $P_1, Q_1, Q_2, \cdots, Q_{p-1}$ of G_{p-1} constitute a set of generators of G_p; that the relations (18) and (18') still hold; and that the only new generating relations of G_p are those which are obtained in the following manner: Let x turn around the origin $n_1 n_2 \cdots n_{p-1}$ times. Then each y_i will be reproduced, and the points y_{0j} will be permuted according to the following cyclic substitution:

$$(y_{00}, y_{0, l_p}, y_{0, 2 l_p}, \cdots),$$

where

$$m_p = h_p n_p + l_p, \qquad 0 \leqq l_p < n_p.$$

There ensues a transformation T_p of the loops g_j, which furnishes the required new generating relations of G_p. It is easily seen that these relations are the following:

$$\begin{aligned} g_{j+l_p} &= A g_j A^{-1}, && (j = 0, 1, \cdots, n_p - l_p - 1), \\ g_j &= A P_p g_{n_p - l_p + j} P_p^{-1} A^{-1}, && (j = 0, 1, \cdots, l_p - 1), \end{aligned} \tag{19}$$

where

$$A = P_1^{m_1 n_2 n_3 \cdots n_{p-1}} P_2^{m_2 n_3 \cdots n_{p-1}} P_{p-1}^{m_{p-1}} P_p^{h_p}.$$

Using (18) and (18′) we can easily show that

$$A = Q_{p-1}^{n_{p-1}} P_p^{h_p}.$$

In fact, from those relations it follows in the first place that P_i *is commutative with* $Q_{i-1}^{n_{i-1}}$. Since this is obviously true for $i = 2$, let us assume that this is true for a given i and let us prove it for $i+1$. Since by hypothesis P_i is commutative with $Q_{i-1}^{n_{i-1}}$, it follows from $Q_i^{n_i} = P_i^{m_i} Q_{i-1}^{n_{i-1} n_i}$ that $Q_i^{n_i}$ is commutative with P_i and with $Q_{i-1}^{n_{i-1}}$. From $P_{i+1} = Q_i^{x_i} P_i^{-y_i} Q_{i-1}^{-n_{i-1} x_i}$ it follows then that $Q_i^{n_i}$ is commutative with P_{i+1}. Using this result, we find:

$$P_1^{m_1 n_2 n_3 \cdots n_{p-1}} P_2^{m_2 n_3 \cdots n_{p-1}} = (Q_1^{n_1 n_2} P_2^{m_2})^{n_3 \cdots n_{p-1}} = Q_2^{n_2 n_3 \cdots n_{p-1}};$$

$$Q_2^{n_2 n_3 \cdots n_{p-1}} P_3^{m_3 n_4 \cdots n_{p-1}} = (Q_2^{n_2 n_3} P_3^{m_3})^{n_4 \cdots n_{p-1}} = Q_3^{n_3 n_4 \cdots n_{p-1}};$$

etc., and finally $A = Q_{p-1}^{n_{p-1}} P_p^{h_p}$.

As in the preceding case the generators g_j, $j > 0$, can be eliminated, and the elimination leads to two new relations between the generators of the reduced set: $P_1, Q_1, Q_2, \cdots, Q_{p-1}$ and g_0. One of these relations is obtained directly from (19) and is the following:

$$g_0 = Q_{p-1}^{n_{p-1} n_p} P_p^{m_p} g_0 (Q_{p-1}^{n_{p-1} n_p} P_p^{m_p})^{-1}. \tag{20}$$

The second relation is obtained by considering instead of the relations (19) the weaker relations:

$$g_{j+1} = B g_j B^{-1}, \ (j = 0, 1, \cdots, n_p - 2), \quad g_0 = B P_p g_{n_p - 1} P_p^{-1} B^{-1}, \tag{21}$$

where

$$B = P_1^{x_p m_1 n_2 n_3 \cdots n_{p-1}} P_2^{x_p m_2 n_3 \cdots n_{p-1}} \cdots P_{p-1}^{x_p m_{p-1}} P_p^{y_p} = Q_{p-1}^{x_p n_{p-1}} P_p^{y_p}.$$

The relations (21) correspond to the transformation $T_p^{x_p}$. It is easily shown that the relations (21) together with the relation (20) lead back to (19). Using the relations (21) in order to express the elements g_j in terms of g_0 and of the remaining generators and substituting into the relation $P_p = g_0 g_1 \cdots g_{n_p - 1}$, we find the required relation:

$$(g_0 P_p^{y_p} Q_{p-1}^{x_p n_{p-1}})^{n_p} = (P_p^{m_p} Q_{p-1}^{n_{p-1} n_p})^{x_p}. \tag{20′}$$

We now put

$$Q_p = (g_0 P_p^{y_p} Q_{p-1}^{x_p n_{p-1}})^{m_p} (P_p^{m_p} Q_{p-1}^{n_{p-1} n_p})^{-y_p}.$$

Then

$$Q_p^{x_p} = g_0 P_p^{y_p} Q_{p-1} \quad , \tag{22}$$

and hence we may take Q_p as a generator to replace g_0. We find then from (20) and (20′):

$$Q_p^{n_p} = g_0^{m_p} Q_{p-1}^{n_{p-1} n_p}, \tag{22′}$$

and conversely (20) and (20′) follow from this relation, if we use (22) as the definition of g_0. The last two relations (22), (22′), together with the generating relations (18), (18′) of G_{p-1}, coincide with the relations (17), (17′), (17″) given in the theorem announced above (where P_{p+1} stays for g_0).

5. *The fundamental homologies of the manifold* $z^k = \Pi(y - y_i)$. Given the Puiseux expansion $y = y(x)$ of a branch Γ_p of genus p, we consider the manifold $z^k = \prod_{i-1}^{n} (y - y_i)$, where n is the order of the branch, and k is an arbitrary positive integer. This k-sheeted cyclic Riemann manifold, which we shall call V, has the given branch Γ_p as locus of branch points. If we limit the range of values of x and y to the boundary of a 4-cell $|x| <$ const., $|y| <$ const., we obtain a closed k-sheeted Riemann manifold M, composed of k samples of the ordinary space, the k sheets being connected along a cylinder whose boundary is the knot associated with the given branch. The topological invariants of V, or of M, are topological invariants of Γ_p. We proceed to derive the fundamental homologies for the one-dimensional cycles on M using the generating relations of the fundamental group G_p of the residual space of Γ_p.

It is known that if permutability relations between the generators of G_p are added to the generating relations, then G_p, being the group of a knot, becomes the infinite free group generated by one element. In the present case this also follows immediately from the fact that G_p is also generated by the loops g_i $(i = 0, 1, 2, \cdots, n-1)$, and these loops are conjugate elements of G_p. Any of these loops can be taken as the generator of the free group. We have seen before that the element P_{p+1} is one of these loops. Hence in the above free group all the elements are powers of P_{p+1}. If we put

$$\nu_i = n_i n_{i+1} \cdots n_p, \qquad (i = 1, 2, \cdots n_p),$$
$$\nu_{p+1} = 1,$$

and if we introduce integers $\lambda_1, \lambda_2, \cdots, \lambda_p$, defined by the following recurrent relations:

$$\lambda_1 = m_1; \quad \lambda_i = m_i + \lambda_{i-1} n_{i-1} n_i,$$

we easily find that in the free group $\{P_{p+1}\}$ the elements P_i and Q_i are represented by the following powers of P_{p+1}:

$$P_i = P_{p+1}^{\nu_i}, \qquad Q_i = P_{p+1}^{\lambda_i \nu_{i+1}}, \qquad (i = 1, 2, \cdots, p).$$

It follows that the elements

$$P_s{}^j = P^s{}_{p+1} P_j P^{-(s+\nu_j)}_{p+1},$$
$$Q_s{}^j = P^s{}_{p+1} Q_j P^{-(s+\lambda_j \nu_{j+1})}_{p+1} \qquad (s = 0, 1, \cdots, k-1,\ j = 1, 2, \cdots, p)$$

produce the identical substitution on the branches of z and hence correspond to 1-cycles on V. It can also be shown that these $2pk$ cycles form a base for the cycles of their dimension.* We next write the fundamental homologies for the 1-cycles on M. In doing this we may neglect all powers of P_{p+1} in which the exponent is a multiple of k, since obviously $P^k{}_{p+1} \sim 0$ on M.

The generating relations (17) and (17′) lead to the following homologies:

$$(23) \qquad \sum_{r=0}^{n_{j-1}} Q^j{}_{i+r\lambda_j \nu_{j+1}} = \sum_{r=0}^{m_{j-1}} P^j{}_{i+r\nu_j} + \sum_{r=0}^{n_{j-1} n_{j-1}} Q^{j-1}_{i+m_j \nu_j + r\lambda_{j-1}\nu_j}.$$
$$(j = 1, 2, \cdots, p;\ i = 0, 1, \cdots, k-1).$$

$$(24) \qquad \sum_{r=0}^{x_{j-1}} Q^j{}_{i+r\lambda_j \nu_{j+1}} = P_i{}^{j+1} + \sum_{r=0}^{y_{j-1}} P^j{}_{i+\nu_{j+1}+r\nu_j} + \sum_{r=0}^{n_{j-1} x_{j-1}} Q^{j-1}_{i+\nu_{j+1}+y_j \nu_j + r\lambda_{j-1}\nu_j},$$

where $P_i{}^{p+1}$ should be replaced by 0.

The relations (23) and (24) constitute a complete system of $2pk$ homologies between the $2pk$ cycles $Q_s{}^j$, $P_s{}^j$. The rank of the matrix of this system is $2pk - R_1$, where R_1 is the Betti number of M. This matrix, which we shall call A, is made up of circulants of order k. Each circulant may be associated with a polynomial of degree $k-1$, whose coefficients are the elements of the first row of the circulant. Let $f_{2\alpha-1,\beta}$ and $f_{2\alpha-1,\beta+p}$ be the polynomials associated with the cycles $P_i{}^\beta$ and $Q_i{}^\beta$ respectively $(i = 1, 2, \cdots, k)$ in the k homologies of (23) in which $j = \alpha (\alpha, \beta = 1, 2, \cdots, p)$. Let $f_{2\alpha,\beta}$ and $f_{2\alpha,\beta+p}$ be the polynomials defined in a similar manner for the system (24). Then

$$f_{2\alpha-1,\alpha} = (x^{m_\alpha \nu_\alpha} - 1)/(x^{\nu_\alpha} - 1),$$
$$f_{2\alpha-1,\alpha+p-1} = x^{m_\alpha \nu_\alpha} (x^{\lambda_{\alpha-1}\nu_\alpha n_{\alpha-1}} - 1)/(x^{\lambda_{\alpha-1}\nu_\alpha} - 1),$$
$$f_{2\alpha-1,\alpha+p} = -(x^{\lambda_\alpha n_\alpha \nu_{\alpha+1}} - 1)/(x^{\lambda_\alpha \nu_{\alpha+1}} - 1);$$
$$f_{2\alpha,\alpha} = x^{\nu_{\alpha+1}} (x^{y_\alpha \nu_\alpha} - 1)/(x^{\nu_\alpha} - 1),$$
$$f_{2\alpha,\alpha+1} = 1,$$
$$f_{2\alpha,\alpha+p-1} = x^{x_\alpha m_\alpha \nu_{\alpha+1}} (x^{\lambda_{\alpha-1} x_\alpha \nu_{\alpha-1}} - 1)/(x^{\lambda_{\alpha-1}\nu_\alpha} - 1),$$
$$f_{2\alpha,\alpha+p} = -(x^{\lambda_\alpha x_\alpha \nu_{\alpha+1}} - 1)/(x^{\lambda_\alpha \nu_{\alpha+1}} - 1),$$

and f_{ij} is identically zero in every other case.

The matrix A is completely described by the matrix $B = || f_{ij} ||$. It can be proved that if $\omega_1, \omega_2, \cdots, \omega_k$ are the k roots of the equation $x^k = 1$, and if $B(\omega_i)$ denotes the matrix obtained from B, by replacing x by ω_i, then A can be brought into a diagonal form, the elements of the main diagonal being the

* See K. Reidemeister, *loc. cit.*

matrices $B(\omega_i)$, while all the other elements are zero-matrices. It follows that the rank of the matrix A is equal to the sum of the ranks of the k matrices $B(\omega_i)$. By induction it can be easily shown that

$$\text{Det. } B = F(x) = f_1(x) f_2(x) \cdots f_p(x),$$

where

$$f_i(x) = (x^{\lambda_i \nu_i} - 1)(x^{\nu_{i+1}} - 1) / (x^{\lambda_i \nu_{i+1}} - 1)(x^{\nu_i} - 1)$$

is a polynomial in x. The proof by induction is based upon the fact that for every $p' < p$ the matrix of the elements of the first p' rows and columns of B is again of the same type as B. Namely, if we write

$$B = B(x, p, n_1, n_2, \cdots, n_p;\ m_1, m_2, \cdots, m_p)$$

in order to indicate the dependency of the matrix B on the characteristic numbers n_i and m_i and on the variable x, we find that its first p' rows and columns form a matrix $B(x^{\nu_{p'+1}}, p', n_1, n_2, \cdots, n_{p'};\ m_1, m_2, \cdots, m_{p'})$. *Hence* $B(\omega_i)$ *is of rank less than* $2p$ *if and only if* ω_i *is a root of the polynomial* $F(x)$. But it can be shown that *if* $B(\omega_i)$ *is of rank* $2p - \sigma$, *then* ω_i *is precisely a* σ*-fold root of* $F(x)$, or, since all the roots of each factor $f_j(x)$ are simple, ω_i is a root of σ of the polynomials $f_j(x)$. Also this statement is proved by induction in the following manner. If $f_j(x)$ is the first of the polynomials $f_1, f_2, \cdots$, which vanishes for $x = \omega_i$, then the last $2p - 2j$ rows and columns of the matrix B form a matrix

$$B(x, 2p - 2j, n_{j+1}, n_{j+2}, \cdots, n_p;\ \lambda_{j+1}, \lambda_{j+2}, \cdots, \lambda_p) = \bar{B},$$

so that

$$\text{Det. } \bar{B} = f_{j+1}(x) f_{j+2}(x) \cdots f_p(x).$$

It is then easily proved that if $B(\omega_i)$ is of rank $2p - \sigma$, the matrix $\bar{B}(\omega_i)$ is of rank $2p - 2j - \sigma + 1$. Assuming that the statement is true for matrices $B(x, p' \cdots)$ where $p' < p$—for $p = 1$ the truth of the statement is obvious—it follows that the statement holds for the given matrix B of order $2p$.

We thus arrive at the following theorem: *The Betti number* R_1 *of the manifold* M *is equal to the number of roots of the polynomial* $F(x)$ (*each root counted to its proper multiplicity*) *which are also roots of the equation* $x^k = 1$.

It is now an easy matter to show that if two singularities are topologically equivalent their associated polynomials $F(x)$ must coincide, and hence the polynomial $F(x)$ is a topological invariant of the singularity. From this it follows immediately that also the genus of the singularity and all its characteristic numbers m_i and n_i are topological invariants.

THE JOHNS HOPKINS UNIVERSITY.

ON A THEOREM OF EDDINGTON.

By Oscar Zariski.

1. In a recent paper,* Eddington considers a set of 4-rowed complex matrices $F_1, F_2, \cdots F_n$, satisfying the following two conditions: 1) $F_i^2 = -E$, where E is the unit matrix; 2) any two matrices of the set are *anticommutative*, i. e., $F_iF_j = -F_jF_i$, if $i \neq j$. He proves that the maximal number of matrices in such a set is *five*, and that if the matrices are conditioned to be either *real* or *pure imaginary*, then *two at most* are real and the other three are pure imaginary.

Eddington's proof is algebraic. It may not be without interest to show that it is possible to derive Eddington's theorem geometrically in a simple and elegant manner by making use of known facts concerning *abelian groups of involutory collineations* in the complex or real projective space. These groups were classified by Study † for projective spaces of any dimensions. In the present case we find that the problem is connected with the so-called *collineation group of a Kummer configuration*, which is a group of order 16 all elements of which, besides the identity, are harmonic biaxial collineations.‡

2. With each matrix F_i of the set, considered as the matrix of coefficients of a linear homogeneous transformation, we may associate a collineation T_i in the complex projective space S_3. The condition $F_i^2 = -E$ says that T_i is a non-degenerate involutory collineation. The condition that any two distinct matrices of the set be anticommutative says that any two collineations T_i and T_j are permutable. Moreover, the n collineations $T_1, T_2, \cdots, T_n$ are all distinct, because if, for instance, T_i and T_j were the same, then the

* On Sets of Anticommutative Matrices," *Journal of the London Mathematical Society*, Vol. 7, Part 1 (January, 1932).

† "Gruppen zweiseitiger Kollineationen," *Göttinger Nachrichten* (1912), pp. 452-479.

‡ In a paper by M. H. A. Newman, "Note on an Algebraic Theorem of Eddington," published in the *Journal of the London Mathematical Society*, Vol. 7, Part 2 (April, 1932), the reader will find a generalization of Eddington's theorem to matrices of any order 2^nq, q odd. Also this more general theorem could be derived, by a method similar to the one given below, from Study's classification of the maximal groups of involutory collineations in S_r, $r = 2^nq - 1$. The number of projectively distinct maximal groups and their structure depends solely on n. It is therefore obvious from a geometric point of view, that the maximal number of anticommutative 2^nq-rowed matrices and 2^n-rowed matrices must be the same. Dr. J. Williamson, who called my attention to Eddington's theorem, was interested in its generalization and arrived at results, which agree with Newman, except for a slight difference in the possible number of real matrices in a set.

elements of the matrix F_i would be proportional to the elements of the matrix F_j, and the two matrices would be commutative, contrary to hypothesis. It follows that the collineations T_i generate an abelian group of involutary collineations, since the product of two permutable involutory collineations is again an involutory collineation. By Study (*loc. cit.*) there are only two distinct maximal groups of involutory collineations in S_3, i. e. such that every group of involutory collineations either coincides with or is a subgroup of one of them: 1) *the tetrahedral group*, 2) *the group of a Kummer configuration*. The elements of the tetrahedral group, which is of order 8, are collineations whose loci of invariant points are opposite elements of a tetrahedron, i. e., either a face and an opposite vertex or two opposite edges. It follows that the collineations T_i cannot belong to the tetrahedral group, because otherwise, assuming this tetrahedron as reference base for coördinates in S_3, the matrices F_i would be reduced simultaneously to the diagonal form and could not be anticommutative.

Having proved that the collineations T_i must belong to the group of a Kummer configuration, we recall that this group of order 16, which we shall denote by G_{16}, is obtained in the following manner. Consider a quadric surface Q in S_3, and the two families of lines on Q. Call the lines of one family-lines u, and the lines of the other family-lines v. Let $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$ be two pairs of lines u harmonically related to each other. Similarly, let $v^1, \overline{v}^1$ and $v^2, \overline{v}^2$ be two harmonically related pairs of lines v. Let U_1 and U_2 be the harmonic biaxial homologies whose axes are the lines $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$ respectively, and let V_1 and V_2 be the harmonic biaxial collineations whose axes are the lines $v^1, \overline{v}^1$ and $v^2, \overline{v}^2$ respectively. The four collineations U_1, U_2, V_1, V_2 generate the group G_{16}. The product $U_1U_2 = U_3$ has as its axes the pair of lines $u^3, \bar{u}^3$, which are harmonically related to both pairs $u^1, \bar{u}^1$ and $u^2, \bar{u}^2$. Similarly the axes $v^3, \overline{v}^3$ of the product $V_3 = V_1V_2$ are harmonically related to both pairs $v^1, \overline{v}^1$ and $v^2, \overline{v}^2$. The axes of the collineation U_iV_j are the diagonals of the skew quadrilateral formed by the lines u^i, $\bar{u}^i$, v^j, $\overline{v}^j$. The collineations of G_{16} leave invariant the quadric Q and transform the lines of each family into the lines of the same family. The collineation U_i $(i = 1, 2, 3)$ leaves each line v invariant and induces in the family of lines u an involution whose double elements are the lines $u^i, \bar{u}^i$. We shall denote this involution by π_i. Similarly let ω_i denote the involution induced in the family of lines v by the collineation V_i.

A last remark about the group G_{16}, which shall be of use in the sequel, is the following. There are 10 quadrics which are left invariant by the G_{16}, one of which is Q, and each of these quadrics is related to the G_{16} in the

same manner as Q. If the collineations of the G_{16} are real, the above quadrics are all real, and namely 9 are real ruled quadrics, while the tenth has no real points (imaginary ellipsoid).*

3. By a proper transformation of the coördinates we first reduce the equation of Q to the following canonical form:

$$x_1x_4 - x_2x_3 = 0,$$

and we also consider the parametric representation of Q:

$$x_1 = u_1v_1, \quad x_2 = u_1v_2, \quad x_3 = u_2v_1, \quad x_4 = u_2v_2,$$

where the parameters u_1, u_2 and v_1, v_2 can obviously be considered as homogeneuos coördinates of a line u and of a line v respectively. We point out immediately that if the matrices F_i are either real or pure imaginary, the collineations T_i are all real, and therefore, by the remark at the end of the preceding section, we may assume Q to be a real ruled quadric. Therefore the transformation of coördinates, which reduces the equation of Q to the above canonical form, is real.

If T is any collineation in G_{16}, and if the equations of the projectivities π and ω induced by T in the system of lines u and in the system of lines v are

$$u'_1 = a_{11}u_1 + a_{12}u_2, \quad u'_2 = a_{21}u_1 + a_{22}u_2,$$

and

$$v'_1 = b_{11}v_1 + b_{12}v_2, \quad v'_2 = b_{21}v_1 + b_{22}v_2,$$

respectively, then the equations of the collineation T are the following:

$$\begin{aligned}
x'_1 &= a_{11}b_{11}x_1 + a_{11}b_{12}x_2 + a_{12}b_{11}x_3 + a_{12}b_{12}x_4,\\
x'_2 &= a_{11}b_{21}x_1 + a_{11}b_{22}x_2 + a_{12}b_{21}x_3 + a_{12}b_{22}x_4,\\
x'_3 &= a_{21}b_{11}x_1 + a_{21}b_{12}x_2 + a_{22}b_{11}x_3 + a_{22}b_{12}x_4,\\
x'_4 &= a_{21}b_{21}x_1 + a_{21}b_{22}x_2 + a_{22}b_{21}x_3 + a_{22}b_{22}x_4.
\end{aligned}$$

The matrix F of the coefficients of T is nothing else than the direct product of the two matrices A and B, in symbols:

$$F = A \times B,$$

where

$$A = \begin{pmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11}, & b_{12} \\ b_{21}, & b_{22} \end{pmatrix}.$$

If the matrix F is given, then the two matrices A and B are determined to within factors α and β connected by the relation: $\alpha\beta = 1$. Moreover, if

* Study, *loc. cit.*, p. 467.

$$F_1 = A_1 \times B_1, \qquad F_2 = A_2 \times B_2,$$

then

$$F_1F_2 = (A_1A_2) \times (B_1B_2). \tag{1}$$

In particular, $F^2 = A^2 \times B^2$. Since both π and ω are involutions, we have $A^2 = \rho E$ and $B^2 = \sigma E$. We shall choose the factor α so as to have always $A^2 = E$. But then, in view of the condition $F^2 = -E$, we must have $B^2 = -E$. A and B are then determined to within a $\pm$ sign.

We have to consider 3 matrices $A : A_1, A_2, A_3$, corresponding to the 3 involutions π_1, π_2, π_3 considered above, and 3 matrices $B : B_1, B_2, B_3$, corresponding to the 3 involutions $\omega_1, \omega_2, \omega_3$. We also introduce the matrices A_0 and B_0, corresponding to the identities π_0 and ω_0. The involutions π_1, π_2, π_3 and the identity π_0 form a group, and if G_{16} is real, then of the 3 involutions π_i, all real, two are necessarily hyperbolic and one is elliptic. We may then suppose that the double elements $u^1, \bar{u}^1$; $u^2, \bar{u}^2$; $u^3, \bar{u}^3$ of the involutions π_1, π_2, π_3 correspond to the following values of the non-homogeneous parameter $u = u_1/u_2 : 0, \infty ; +1, -1 ; +i, -i$. In view of the condition $A_i^2 = E$, we find then:

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Making a similar supposition for the double elements of the involutions $\omega_1, \omega_2, \omega_3$ we find in view of the condition $B_i^2 = -E$, that

$$B_0 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \quad B_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad B_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Each of the matrices $F_1, F_2, \cdots, F_n$ is the direct product of an A_i by a B_j. Let

$$F_k = A_{i_k} \times B_{j_k}, \qquad (k = 1, 2, \cdots, n)$$

where i_k and j_k are numbers of the set 0, 1, 2, 3. We can verify directly that $A_iA_j = -A_jA_i$ and $B_iB_j = -B_jB_i$ if $i \neq j$ and neither i nor j has the value 0. In the excluded cases the matrices will be of course commutative. We now use the condition that $F_kF_l = -F_lF_k$. We deduce from this condition and from (1) that the matrices A_{i_k} and A_{i_l} are commutative or anticommutative according as the matrices B_{j_k} and B_{j_l} are anticommutative or commutative. I say that if no i_k *and* no j_k is 0, then $n \leqq 3$. In fact, if all the i_k's have the same value, then any two of the matrices A_{i_k} are commutative, and therefore any two of the matrices B_{j_k} must be anticommutative, and that

is possible only if all the j_k's are distinct and different from 0, i. e., if $n \leqq 3$. If not all the i_k's are equal, let, for instance, $i_1 \neq i_2$. Then necessarily $j_1 = j_2$. For any other value of k we must also have $j_k = j_1$. In fact, if $j_k \neq j_1$, then B_{j_k} is anticommutative with $B_{j_1} = B_{j_2}$ and therefore A_{i_k} must be commutative with both A_{i_1} and A_{i_2}, which is impossible, since $i_1 \neq i_2$ and $i_k \neq 0$. Hence all the j_k's have the same value, and we conclude, as before, that $n \leqq 3$. Hence, if $n > 3$, then for some value of k we must have either $i_k = 0$ or $j_k = 0$. Suppose that, for instance, $i_1 = 0$. Then obviously $j_k \neq 0$ for any k. If now $n > 4$, then two indices i_k must be equal. If these two equal indices are not 0, then let, for instance, $i_2 = i_3 = 1$. Since any two of the matrices A_0, A_1, A_1 are commutative, the indices j_1, j_2, j_3 must be all distinct and different from 0. Let, for instance, $j_1 = 1$, $j_2 = 2$, $j_3 = 3$. For any other value of $k > 3$, we must have $j_k \neq 1$ and also $j_k \neq 2, 3$, since B_{j_k} must be either commutative or anticommutative with both B_{j_2} and B_{j_3}. Hence $j_k = 0$, which is impossible. Consequently, if $n > 4$, the two equal indices i_k must be equal to 0. Let $i_1 = i_2 = 0$. Then j_1 and j_2 must be distinct and again no j_k can be equal to 0. For any $k > 2$, we must have $j_k \neq j_1, j_2$. Hence $j_3 = j_4 = j_5 = \cdots = j_n$, and for this reason the indices $i_3, i_4, i_5, \cdots, i_n$ must be all distinct and different from 0. We conclude that $n = 5$ and, for instance, that $i_3 = 1$, $i_4 = 2$, $i_5 = 3$, and that the matrices F_k are the following:

$$F_1 = A_0 \times B_{j_1},\ F_2 = A_0 \times B_{j_2},\ F_3 = A_1 \times B_{j_3},\ F_4 = A_2 \times B_{j_3},\ F_5 = A_3 \times B_{j_3},$$

where the numbers j_1, j_2, j_3 coincide with the numbers 1, 2, 3, in some order. This proves the first part of Eddington's theorem.

From the explicit expression of the matrices A_i and B_i given above, we see that if $j_3 = 3$ then the matrices F_3 and F_4 are real and the others are pure imaginary, whereas if $j_3 \neq 3$, say if $j_1 = 1$, $j_2 = 3$, $j_3 = 2$, then F_2 and F_5 are real and the others are pure imaginary. This proves the second part of Eddington's theorem.

THE JOHNS HOPKINS UNIVERSITY.

REGULAR LINEAR SYSTEMS OF CURVES WITH THE SINGULARITIES OF A GIVEN CURVE AS BASE POINTS.

By MARGUERITE LEHR.

Summary. An irreducible algebraic curve of order $m, f(x, y) = 0$, in the plane $\pi(z = 0)$ is given with singularities of a specified kind. The surface $F, z^n = f(x, y)$ where n is a positive integer, gives rise to adjoint surfaces ϕ_ν, and to complete linear systems of surfaces $\phi_\nu - d\pi$ which together with the plane π taken d times form adjoints ϕ_ν. These systems of surfaces cut the plane π in systems of curves $| C^d |$, shown to be complete. The irregularity of the surface F can be expressed in terms of the superabundances of some of the systems $| C^d |$, hence application of the theorem * that F is regular if n is a power of a prime shows the regularity of the systems of plane curves involved. For $f(x, y) = 0$ having only nodes and cusps, Zariski † obtains in this manner regular systems of curves with the cusps of f as base points. In this paper we consider f possessing either of two types of singularities:— ordinary k-fold points and k-fold points each of which is the origin of one branch of order k. The resulting theorems are analogous to the well-known theorem: Curves of order $m - 3$ with $(k - 1)$-fold points at the k-fold points of f form a regular system; we prove that curves of order $m - 3 - j$ with $(k - i)$-fold points $(i = 1, 2, \cdots, k - 1)$ at the k-fold points of f form regular systems for j having a certain range of values. Thus restrictions are imposed on the number and position of specified singularities for f. Further, since the surface F is regular for all positive integral values of n when $f(x, y)$ has a cyclic fundamental group with respect to its carrying complex plane,† the connection between the irregularity of F and the superabundances of the linear systems described enables us to state sufficient conditions that $f(x, y) = 0$ have a non-cyclic fundamental group; these conditions show that there exists a connection between the position of the singular points other than nodes and the structure of the fundamental group.

The adjoints $\phi_\nu(x, y, z) = 0$ are conditioned in general at all singularities of $z^n = f(x, y)$. This surface F has isolated singularities only at the singular

* O. Zariski, "On the Linear Connection Index of the Algebraic Surfaces $z^n = f(x, y)$," *Proceedings of the National Academy of Sciences*, Vol. 15 (No. 6, June, 1929).

† O. Zariski, "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

points of f when $n = m$; if $n > m$ the line at infinity in the plane π is multiple on F, the only hypermultiple points on it being intersections with $f(x, y) = 0$. This line may be considered in generic position with respect to f, having m distinct intersections with f. In the paper quoted,* Zariski gives a necessary and sufficient condition imposed on adjoints ϕ_ν by the multiple line, and shows that the hypermultiple points on it impose no additional conditions. We shall determine the conditions imposed on ϕ_ν at isolated singularities of F due to various types of singular points on $f(x, y) = 0$. These together with the condition mentioned will completely describe the adjoint surfaces ϕ_ν. For the sake of clarity, the case of f possessing only ordinary k-fold points will be handled fully first; the procedure can then be modified for other conditions on f.

I. f HAVING ORDINARY k-FOLD POINTS ONLY.

1. *Necessary and sufficient conditions on ϕ_ν at O.* Let O be an ordinary k-fold point of $f(x, y) = 0$, which we may suppose to be at the origin. If $\phi_\nu(x, y, z) = 0$ is to be adjoint to $z^n = f(x, y)$, the double integral

$$\int\int \frac{\phi_\nu(x, y, z)}{z^{n-1}}\, dxdy$$

must be finite for every analytical 2-cell containing $x = y = 0$. Since the required condition is of differential character, we may replace $f(x, y)$ by the product $\Pi(y - a_i x) = (y - a_1 x)(y - a_2 x) \cdots (y - a_k x)$. Introduce two new independent variables u, t by the relations

$$x = u^n, \quad y = u^n t, \quad z^n = u^{kn}\Pi(t - a_i) = u^{kn}\Pi(t). \tag{1}$$

Let $c_{\alpha\beta\gamma}x^\alpha y^\beta z^\gamma$ be the general term of the polynomial $\phi_\nu(x, y, z)$. The transformed double integral is then the sum of integrals of the form

$$\int\int \frac{u^{\alpha n+\beta n+\gamma k+2n-1} t^\beta \{P(t)\}^{\gamma/n}}{u^{k(n-1)}\{P(t)\}^{n-1/n}}\, dudt,$$

which must be finite for $u = 0$, t arbitrary. So far as t is concerned, the integral is finite for all finite values of t including those for which $P(t) = 0$; the infinite value for t can be investigated by using $x = u^n t, y = u^n$, which leads to the same form of integral. Hence a *necessary and sufficient condition that the double integral remain finite is*

$$(\alpha + \beta - k + 2)n + k(\gamma + 1) > 0. \tag{2}$$

This condition requires $\alpha + \beta + \gamma \geqq k - 2$ and is satisfied by

* O. Zariski, "On the Irregularity of Cyclic Multiple Planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

$$\alpha + \beta = k - 2, \ \gamma = 0;$$

so every ϕ_ν has at least a $(k-2)$-fold point at O, and the generic ϕ_ν cuts the plane $\pi(z=0)$ in a curve C of order $\nu - \sigma_0$ (where σ_0 accounts for the line at infinity occurring in the intersection), having an ordinary $(k-2)$-fold point at O. On the other hand, consider any curve of order $\nu - \sigma_0$ with the points * $O(k-2)$-fold in the plane π, and project it from the point at infinity on the z-axis. The resulting cone satisfies the adjoint condition at O given by (2). Since y^{σ_0} satisfies the conditions imposed † by the multiple line at infinity in the plane π, the degenerate surface made up of y^{σ_0} and the cone is an adjoint ϕ_ν. Every such curve can be cut out thus by an adjoint surface, so the system of curves cut out on π by $|\phi_\nu|$ is complete.

2. *Nature of the curves* $|C^d|$. Consider all surfaces $\phi_{\nu-d}$ which with the plane π taken d times give ϕ_ν adjoint. They form a complete linear system which we denote by $|\phi_\nu - d\pi|$, and this system cuts the plane π in a system of curves $|C^d|$. We wish to describe these curves for arbitrary values of d. Surfaces $\phi_\nu - d\pi$ are those ϕ_ν satisfying (2), in whose terms $\gamma \geq d$. When they are cut by $z=0$, the curves C^d are given by terms in ϕ_ν in which $\gamma = d$; hence in the equation of $C(x,y)=0$, the exponents of $x^\alpha y^\beta$ satisfy the relation

$$(\alpha + \beta - k + 2)n + k(d+1) > 0. \tag{3}$$

The surfaces ϕ_ν cut π in a $(k-2)$-fold point at O, and we know that there are in addition, in general, infinitely near multiple points to O, so removing $z=0$ as a factor from ϕ_ν will not necessarily reduce the multiplicity of O on the curve of section. From (3) we see that if terms $\alpha + \beta = k - 3$ are to be present in the equation of C, then

$$d \geq [n/k]$$

where $[n/k]$ means the integral part of n/k. So the systems

$$|\phi_\nu|, \ |\phi_\nu - \pi|, \ \cdots \ |\phi_\nu - \{[n/k] - 1\}\pi|$$

cut $z=0$ *in curves with* $(k-2)$*-fold points at the* k*-fold points of* f. These points are ordinary $(k-2)$-fold points since any terms for which

$$\alpha + \beta = k - 2$$

satisfy condition (3). The curves C^d cut out by $|\phi_\nu - d\pi|$ can and will

* The only singularities of f are ordinary multiple points of one order k; we shall say "the points O" meaning that set of points.

† O. Zariski, "On the irregularity, etc.," *loc. cit.*

have O as a point of multiplicity $k-2-i$ if and only if d satisfies the relation

$$-in+kd+k>0 \text{ or } d\geq[in/k]. \qquad (i=0,1,\cdots,k-2).$$

The curves C^d, then, cut out on π by $|\phi_\nu - d\pi|$ have the ordinary k-fold points of f as ordinary $(k-2-i)$-fold points for values of d given by

$$(4) \qquad [in/k]\leq d<[(i+1)n/k]. \qquad (i=0,1,\cdots,k-2).$$

Finally for $d\geq[(k-2)n/k]$, the systems $|\phi_\nu - d\pi|$ are unconditioned at the ordinary k-fold points of $f(x,y)=0$.

The order of the curves $|C^d|$ is $\nu-d-\sigma_d$, where σ_d accounts for the multiple line at infinity. Write

$$(5) \qquad n-m=hm-m_1, \qquad h>0,\ 0\leq m_1<m.$$

For values $d\leq n-2$, $\nu-d-\sigma_d=\nu-n+m+1-\mu$, where μ is the positive integer defined * by

$$(6) \quad (h+1)\mu+[\mu m_1/m]-2<d\leq(h+1)(\mu+1)-[(\mu+1)m_1/m]-2.$$

For $d\geq n-2$, the surfaces are not only unconditioned at the multiple line, but they are also unconditioned at the k-fold points of $f(x,y)=0$, since $[(k-2)n/k]\leq n-2$ for all positive integral k and n. For $d\geq n-2$, the order of $|C^d|$ is $\nu-d$. *The systems*

$$|C^0|,\ |C^1|,\cdots,\ |C^{n-2}|$$

therefore distribute themselves into sets of consecutive systems all of order $\nu-n+m+1-\mu$ where μ runs from 0 to $m-1$. There are $h+\epsilon_\mu$ systems in each set, where

$$(7) \qquad \begin{aligned} &\epsilon_\mu=1 \text{ if } [\mu m_1/m]=[(\mu+1)m_1/m] \text{ and } \mu\neq 0,\\ &\epsilon_\mu=0 \text{ if } [\mu m_1/m]=[(\mu+1)m_1/m]-1, \text{ or } \mu=0. \end{aligned}$$

In particular, the first h systems are of order $\nu-n+m+1$, and the last h of order $\nu-n+2$.

3. *Completeness of* $|C^d|$. Consider any curve of order $\nu-d-\sigma_d$ lying in the plane π and having points of multiplicity

$$k-2-i \qquad (i=0,1,\cdots,k-2)$$

at an ordinary k-fold point of f. The cone obtained by projecting this curve

* O. Zariski, "On the irregularity of cyclic multiple planes," *Annals of Mathematics*, 2nd ser., Vol. 32 (No. 3, June, 1931).

from the point at infinity on the z-axis together with the plane π taken d times satisfies condition (2); it is a surface, therefore, which satisfies the conditions imposed on adjoints at isolated singularities of F. As previously stated, $y^{\sigma d}$ will satisfy the conditions imposed by the multiple line, hence the cone together with σ_d times y and d times π will give an adjoint ϕ_ν. The cone plus σ_d times y is then a $\phi_\nu - d\pi$, and since every C^d of the type described can be cut out by such a surface, the systems $| C^d |$ cut out by $| \phi_\nu - d\pi |$ and described by the relations (4) and (6) are complete.

4. *Virtual and effective dimensions of* $| \phi_\nu - d\pi |$. If a system $| \phi_\lambda |$, linear and complete, cuts a plane π in a complete system of curves $C_{\lambda-\sigma}$, and if we denote by $| \phi_\lambda - \pi |$ the complete system of surfaces $\phi_{\lambda-1}$ which with π make a ϕ_λ, then for π a generic plane it is well-known that

$$(r_\lambda - r'_\lambda) = (r_{\lambda-1} - r'_{\lambda-1}) + s_\lambda,$$

where r_λ, r'_λ are respectively the effective and virtual dimensions of $| \phi_\lambda |$; $r_{\lambda-1}$, $r'_{\lambda-1}$ respectively those of $| \phi_\lambda - \pi |$, and s_λ is the superabundance of the complete system of curves C_λ cut out on π (σ being zero). In the case of the systems $| \phi_\nu - d\pi |$, however, the plane is not generic; it goes through the multiple line and all isolated singularities. It is true, nevertheless, that $| \phi_\nu - d\pi |$ cuts π in a complete system $| C^d |$, and we shall show that under this condition

$$r_d - r'_d = r_{d+1} - r'_{d+1} + s_d$$

where r_d, r'_d are respectively the effective and virtual dimensions of $| \phi_\nu - d\pi |$ and s_d is the superabundance of C^d. If we denote by ρ_d, ρ'_d the effective and virtual dimensions respectively of the complete system $| C^d |$, we can write

$$r_d - r'_{d+1} = \rho_d + 1.$$

Since

$$r'_d = \binom{\nu - d + 3}{3} - 1 - K_d(\nu - d) - K'_d,$$

$$r'_{d+1} = \binom{\nu - d + 2}{3} - 1 - K_{d+1}(\nu - d) - K'_{d+1},$$

K_d, K'_d, K_{d+1}, K'_{d+1} being the postulation constants independent of ν, for any fixed d, or, in other words, algebraic functions of d alone, we obtain

$$(8) \qquad r'_d - r'_{d+1} = \binom{\nu - d + 2}{2} - (\nu - d)(K_d - K_{d+1}) - (K'_d - K'_{d+1})$$

valid for all non-negative integral values of ν and of d. Now when $\nu - d$ is sufficiently high, $r_d = r'_d$ and $r_{d+1} = r'_{d+1}$; it follows that

$$r'_d - r'_{d+1} = r_d - r_{d+1} = \binom{\nu - d - \sigma_d + 2}{2} - K_d'' + s_d$$

where K_d'' is the postulation in the plane on the curves C^d. Since σ_d is the multiplicity of the line at infinity in π on $\phi_\nu - d\pi$, σ_d is independent of ν; for high values of ν the order of the curves C^d increases and s_d becomes zero. For $\nu - d$ sufficiently high, then,

$$r'_d - r'_{d+1} = \rho'_d + 1. \tag{9}$$

The two expressions in (8), (9) are equivalent for all non-negative integral values of ν and of d for which $\nu - d$ is sufficiently high, *i. e.*,

$$\binom{\nu - d + 2}{2} - (\nu - d)(K_d - K_{d+1}) - (K'_d - K'_{d+1}) = \binom{\nu - d - \sigma_d + 2}{2} - K_d'' + s_d,$$

the K's and σ_d being independent of ν. Fix d at a value d_0. The two polynomials in ν are identical for all ν sufficiently high, hence they are identical polynomials and their coefficients (algebraic functions of d_0) are the same. These relations between the coefficients hold for all non-negative integral values of d, hence the coefficients are identical and the two expressions are equivalent for all values of ν, d. Consequently, so long as $|\phi_\nu - d\pi|$ cuts π in a complete system

$$r'_d - r'_{d+1} = \rho'_d + 1$$

regardless of whether π is generic or not. Since

$$r_d - r_{d+1} = \rho'_d + s_d + 1,$$

we obtain the desired relation

$$r_d - r'_d = r_{d+1} - r'_{d+1} + s_d. \tag{10}$$

5. *Irregularity of F in terms of s_d.* Adding the relations (10) for $d = 0, 1, \cdots, [(k-2)n/k] - 1$, we find

$$r_0 - r'_0 = s_0 + s_1 + \cdots + s_{[(k-2)n/k]-1} + (r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}). \tag{11}$$

This relation holds even if the systems $|\phi_\nu - d\pi|$ or $|C^d|$ cease to exist, if we adopt the convention that the effective dimension of a non-existent system is -1.

For $\nu \geq n - 4$, $r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}$ can be shown to be zero. Consider the subadjoint surfaces $\bar{\phi}_\nu$; they are conditioned at the multiple line as are ϕ_ν, and are not conditioned at the isolated singularities, so

$$| \phi_\nu - [(k-2)n/k]\pi | \text{ and } | \bar{\phi}_\nu - [(k-2)n/k]\pi |$$

are coincident systems. Now all systems $| \bar{\phi}_\nu - d\pi |$ cut out complete systems $| \bar{C}^d |$ by even stronger reasoning than for $| \phi_\nu - d\pi |$, so

$$\bar{r}_d - \bar{r}'_d = \bar{r}_{d+1} - \bar{r}'_{d+1} + \bar{s}_d.$$

But $| \bar{C}^d |$ is virtually and effectively the system all curves of order $\nu - d - \sigma_d$, of the same order as $| C^d |$, but unconditioned at isolated singularities, so $\bar{s}_d = 0$, if the order is ≥ -2. It follows that

$$\bar{r}_0 - \bar{r}'_0 = \bar{r}_{d+1} - \bar{r}'_{d+1},$$

for any d such that $\nu - d - \sigma_d \geq -2$. When $d \leq [(k-2)n/k] - 1$, the order of C^d is $\geq \nu - n + 2$, so if $\nu \geq n - 4$, curves C^d are all of order greater than or equal to -2. When

$$d = [(k-2)n/k] - 1,$$

then

$$\bar{r}_0 - \bar{r}'_0 = r_{[(k-2)n/k]} - r'_{[(k-2)n/k]}.$$

Since $| \phi_\nu - [(k-2)n/k]\pi |$ is identical with $| \bar{\phi}_\nu - [(k-2)n/k]\pi |$. Now by a theorem of Castelnuovo,* $\bar{r}_0 - \bar{r}'_0$ will be zero if $| \phi_{\nu+1} |, | \phi_{\nu+2} |, \cdots$ cut out on a generic plane complete and regular linear systems of curves. If $\nu \geq n - 4$, these systems cut out curves of order $\geq n - 3$, adjoint to the plane section of F, and it is well-known that these curves form regular systems. They are clearly complete, for any curve $C_{\nu'}$ in a generic plane ω, adjoint to the section of F made by ω, can be projected from the point at infinity on the x-axis giving a cone which is a subadjoint surface $\bar{\phi}_{\nu'}$ of F. This establishes the fact that $r_{[(k-2)n/k]} - r'_{[(k-2)n/k]} = 0$ if $\nu \geq n - 4$. Restating (11), then, for $\nu = n - 4$, we obtain the following expression for the irregularity of F:

$$q = s_0 + s_1 + \cdots + s_{[(k-2)n/k]-1}. \tag{12}$$

The systems $| C^d |$ involved all occur in the set $| C^0 |, | C^1 |, \cdots, | C^{n-2} |$ described in § 2, and s_d is the superabundance of the system $| C^d |$ for which the order is given by (6) and (7), and the multiplicity at the ordinary k-fold points of f is given by (4). Hence *the irregularity q of F is equal to the sum of the superabundances of the first $[(k-2)n/k]$ systems $| C^d |$ described.*

6. *Regular systems* $| C^d |$. If f is irreducible, curves of order $m - 3$

* G. Castelnuovo, "Alcune proprietà fondamentali dei sistemi lineari di curve tracciate ecc . . . ," *Annali di Matematica pura ed applicata*, ser. 2, Vol. 25 (1897).

with k-fold points of f as $(k-1)$-fold points form a regular system, so s_0 and possibly certain others succeeding it are zero. Suppose that $m-3-j$ is the maximum order for which the base points O do not give rise to a regular system, *i. e.*, curves of order $m-3-\mu$ with points $O(k-2)$-fold are regular and $s_\mu{}^{k-2}$ is zero if $\mu < j$, but $s_\mu{}^{k-2} > 0$ for $\mu = j$. The irregularity of F involves such superabundances in the first $[n/k]$ terms, hence if the order $m-3-j$ is such that the number of systems preceding it in the sequence is less than $[n/k]$, a non-vanishing superabundance will appear in (12) and the surface F will be irregular. The first system of order $m-3-j$ is preceded by

$$\sum_{i=0}^{j-1} h + \epsilon_i = j(h+1) - 1 - [jm_1/m]$$

terms. Denote the difference between this and $[n/k]$ by λ.

$$\lambda = j(h+1) - 1 - [jm_1/m] - [n/k]. \tag{13}$$

Then if $\lambda < 0$, $q > 0$ and F is irregular. Write

$$n = \rho k + \sigma, \qquad \rho > 0,\ 0 \leq \sigma < k.$$

Recalling that $n = (h+1)m - m_1$, we find from (5) and (13)

$$k\lambda = (h+1)(kj - m) - k - k[jm_1/m] + m_1 + \sigma. \tag{14}$$

As n becomes indefinitely large, h increases indefinitely, but all other elements in λ are either fixed or have an upper limit set by k or by m, so λ takes its sign from $kj - m$. If $kj < m$, λ would be negative for all values of n sufficiently high. We know, however, that if n is a power of a prime, F is regular and λ cannot be negative for such values. Hence kj must be greater than or equal to m. We state the geometric result: If $f(x, y) = 0$ be irreducible of order m, with ordinary k-fold points O, and if μ be the maximum integer such that $k\mu < m$, then the complete systems of curves of orders $m-3$, $m-4, \cdots, m-3-\mu$ with the points $O(k-2)$-fold are regular.

Next, let $m-3-j$ be the maximum order for which the points $O(k-3)$-fold form the base of a superabundant system. For each system $|C^d|$, $d = 0, 1, \cdots, [2n/k]-1$, set up a system of the same order with $O(k-3)$-fold; the superabundance $s_d{}^{k-3}$ of such a system is not greater than the superabundance s_d occurring in q. If in the sequence $|C^0|, \cdots, |C^{n-2}|$, then, the first system of order $m-3-j$ is preceded by less than $[2n/k]$ others, a non-vanishing superabundance will occur in (12) and F will be irregular. Writing

$$\lambda = j(h+1) - 1 - [jm_1/m] - [2n/k],$$

we find, precisely as above, $kj \geq 2m$. Hence for $k\mu < 2m$, the systems of order $m - 3 - \mu$ with points $O(k-3)$-fold are regular. The general theorem may be derived by application of this method, using (4) to give the number of systems $| C^d |$ involved when d is set.

THEOREM. *If $f(x, y) = 0$ be irreducible of order m, with only ordinary k-fold points O, and if μ be the maximum integer such that $k\mu < (i+1)m$, $(i = 0, 1, \cdots, k-2)$ then all complete systems of curves of orders $m-3$, $m-4, \cdots, m-3-\mu$ with the points $O(k-2-i)$-fold are regular.*

7. *Conditions on the k-fold points of $f(x, y)$.* The theorem just proved excludes the possibility of an irreducible curve having so many k-fold points that the virtual dimension of any of the systems described becomes less than -1. It also precludes the ordinary k-fold points of an irreducible curve from having such special positions in the plane that they do not impose independent conditions on the curves of the orders stated, when used to the multiplicity described in the base set. An example will show how this condition operates. A curve of order 14 can have 26 triple points; by the theorem just proved, curves of order 7 passing simply through these points must form a regular system. (Note that for f irreducible, a repeated application of the theorem on adjoints of order $m-3$ will give curves of order eight through the triple points forming a regular system; the present theorem lowers the degree here.) Fix sixteen of the eighteen intersections of a cubic and a sextic. Any C_7 through the fixed points cuts the sextic in addition in 26 points which cannot be triple points of an irreducible C_{14}, as the following argument shows. Consider all C_7's through the 26 points; the characteristic series cut out on a generic C_7 of the system is a complete g_{23}, and the system is regular or superabundant according as the series is non-special or special.* Among the groups of the g_{23} is one made up of the sixteen points fixed on the sextic and an arbitrary line section of C_7. This group is also cut out on C_7 by a quartic made up of the given cubic (with its additional five points of intersection with C_7 fixed) and an arbitrary line, hence this group is contained in the complete series cut out by canonical adjoints through five fixed points. Since the genus of a non-singular C_7 is 15, the canonical series is g_{28}^{14}; the complete series determined by the group of 23 points is therefore g_{23}^{9}, special; hence the system of curves C_7 is superabundant. This contradicts the theorem; the group of 26 points described cannot be triple points of an irreducible C_{14}.

---

* Guido Castelnuovo, "Ricerche generali sopra i sistemi lineari di curve piane," *Torino Memorie*, Ser. 2, Vol. 42 (1891).

II. f HAVING AT EACH k-FOLD POINT ONE BRANCH OF ORDER k.

8. *Necessary and sufficient condition on ϕ_ν at O.* Again suppose O, k-fold, to be at the origin, and consider first O as the origin of a single branch of order k and class 1. The condition being of differential character, $f(x, y)$ may be replaced by $y^k + x^{k+1}$. If $\phi_\nu(x, y, z)$ is to be adjoint to $z^n = f(x, y)$, the double integral

$$\int\int \frac{\phi_\nu(x, y, z)}{z^{n-1}}\, dxdy$$

must be finite for every analytical two-cell containing $x = y = 0$. Introduce two new variables u, t by the relations:

$$x = u^{nk},\ y = u^{n(k+1)}t.$$

By considerations precisely like those of part I, we find that a *necessary and sufficient condition that the double integral remain finite is*

$$(15) \qquad kn(\alpha + \beta - k + 1) + n(\beta + 1) + k(k + 1)(\gamma + 1) > 0.$$

Clearly, if $\alpha + \beta \geq k - 1$, the condition is satisfied; further, if $n \geq k(k + 1)$ no terms for which $\alpha + \beta + \gamma < k - 1$ can appear. So the generic ϕ_ν cuts the plane π in a curve with O as $(k - 1)$-fold point. As in part I, we see that any curve in π with O as ordinary $(k - 1)$-fold point can be projected from the point at infinity on the z-axis, giving a cone which satisfies the conditions imposed on adjoints by the point O, and the system cut out on π by $| \phi_\nu |$ is complete.

9. *Nature of the curves* $| C^d |$. We consider the systems of curves $| C^d |$ cut out on π by the systems of surfaces $| \phi_\nu - d\pi |$. The curves $| C^d |$ are given in the plane $z = 0$ by terms $x^\alpha y^\beta z^d$ of ϕ_ν, hence terms in the equation of $C(x, y) = \sum c_{\alpha\beta} x^\alpha y^\beta$ satisfy the relation

$$(16) \qquad kn(\alpha + \beta - k + 1) + n(\beta + 1) + (d + 1)k(k + 1) > 0.$$

The system $| C^0 |$ has points $O(k - 1)$-fold with no further restriction. If terms $\alpha + \beta = k - 2$ are to satisfy the condition, the minimum value of d is given by the maximum $\beta = k - 2$, or

$$d \geq [n/k(k + 1)].$$

Hence the systems $| \phi_\nu |, | \phi_\nu - \pi |, \cdots, | \phi_\nu - \{[n/k(k + 1)] - 1\}\pi |$ cut the plane in curves with points O as $(k - 1)$-fold points. Stating the general case, then, the curves cut out by $| \phi_\nu - d\pi |$ can and will have O as multiple point of order $k - 1 - i$ if and only if d satisfies the relation

$$(17) \qquad -ikn + n(\beta+1) + k(k+1)(d+1) > 0, \qquad (i = 0, 1, \cdots, k-1).$$

Again the minimum d satisfying this condition is given by the maximum $\beta = k-1-i$, so

$$d \geq [\{(i-1)k + i\}n/k(k+1)].$$

The curves C^d cut out by $|\phi_\nu - d\pi|$ have O as $(k-1-i)$-fold point for values of d given by

$$(18) \qquad \left[\frac{(i-1)k+i}{k(k+1)}\, n\right] \leq d < \left[\frac{(ik+i+1)}{k(k+1)}\, n\right].$$

In particular, for $d \geq [(k^2-k-1)n/k(k+1)]$, the curves are unconditioned at points O.

In addition, the singularity at O on $|C^d|$ will in general be specialized, having a certain number of tangents coinciding with the tangent to f at O. Actually the condition (17) gives

$$(19) \qquad d \geq \left[\frac{(ik-1-\beta)n}{k(k+1)}\right]$$

for $O(k-1-i)$-fold or less. The maximum value of β gives the minimum value of d stated in (18); for this, $|C^d|$ has all tangents at O coincident with the tangent to f at O. As β decreases, d definitely increases until we reach $d = [(ik-1)n/k(k+1)]$ which gives $O(k-1-i)$-fold with all tangents freed from coincidence with the tangent to f at O. From this value to $d = [(ik+i+1)n/k(k+1)]$ the curves $|C^d|$ have the same behavior at O. The order of the curves is given in part I. For n sufficiently large, i. e., if

$$[(k^2-k-1)n/k(k+1)] \leq n-2$$

the surfaces $|\phi_\nu - d\pi|$ for $d \geq n-2$ are unconditioned at the isolated points as before. Hence *the systems $|C^0|, \cdots, |C^{n-2}|$ distribute themselves into sets of consecutive systems all of orders $\nu - n + m + 1 - \mu$, where μ runs from 0 to $m-1$. Any given $|C^d|$ has its multiplicity $k-1-i$ at points O determined by* (18), *and it has at each such point τ tangents coinciding with the tangent to f there, where τ is the minimum integer satisfying the relation*

$$-ikn + n(\tau+1) + (d+1)k(k+1) > 0.$$

10. *Irregularity of F in terms of s_d.* The systems $|C^d|$ are complete, by considerations analogous to those in part I, hence the results of paragraph 4 there apply; further, for n sufficiently high, the surfaces $|\phi_\nu - d\pi|$ for $d = [(k^2-k-1)n/k(k+1)] \equiv \delta$ are conditioned at the multiple line

only, so the argument of paragraph 5 on subadjoint surfaces can be applied. Summing (10) then for values $d = 0, 1, \cdots, [(k^2 - k - 1)n/k(k+1)]$, we obtain

$$r_0 - r'_0 = s_0 + s_1 + \cdots + s_{\delta-1} + r_\delta - r'_\delta. \tag{20}$$

For $\nu \geq n - 4$, therefore,

$$q = s_0 + s_1 + \cdots + s_{\delta-1} \tag{21}$$

where the s_d are superabundances of the systems $| C^d |$ described in 9.

11. *Regular Systems.* If f is irreducible, curves of order $m - 3$ with the points O as $(k-1)$-fold points form a regular system, so s_0 and possibly others are zero. Suppose that $m - 3 - j$ is the maximum order for which the base points $O(k-1)$-fold give rise to a superabundant system. The first system of order $m - 3 - j$ in the set $| C^0 | \cdots | C^{n-2} |$ is preceded by $j(h+1) - 1 - [jm_1/m]$ systems. If this is less than $[n/k(k+1)]$, a non-vanishing superabundance will occur in (21) and F will be irregular. If $k(k+1)j < m$, then, F will be irregular for all n sufficiently high, as in part I, contradicting the established theorem that $z^n = f(x, y)$ is regular for n a power of a prime. We deduce that $k(k+1)j \geq m$; *for any μ satisfying $k(k+1)\mu < m$, curves of order $m - 3 - \mu$ with the points $O(k-1)$-fold are regular.*

To state the general case, let $m - 3 - j$ be the maximum order for which curves with the points $O(k - 1 - i)$-fold do not form a regular system. We find that

$$k(k+1)j \geq (ik + i + 1)m.$$

The geometric result may be stated thus: *If $f(x, y) = 0$ be irreducible of order m, with singular points O, each of which is the origin of one branch of order k and class* 1, *and if μ be the maximum integer such that $k(k+1)\mu < (ik + i + 1)m$, then all complete systems of curves of orders $m - 3$, $m - 4$, $\cdots$, $m - 3 - \mu$ with the points $O(k - 1 - i)$-fold are regular.* Similar theorems can be stated for curves having the points $O(k - 1 - i)$-fold, and having in addition at each point O a given number τ of tangents coinciding with the tangent to f at that point, using in each case the value for d given in (19). In particular, if μ is the maximum integer satisfying $k(k+1)\mu < (k^2 - k - 1)m$, then curves of order $m - 3 - \mu$ passing simply through the points O form a regular system; if μ is the maximum integer given by $k(k+1)\mu < (k^2 - 2k - 2)m$, curves of order $m - 3 - \mu$ passing through the points O in contact with f at each point O form a regular system.

12. *Conditions on such singularities of f.* The theorem just stated

imposes certain conditions on the number and position of singularities O possible on an irreducible curve of given order. The question of existence of an irreducible curve possessing a stated set of singularities can in some cases be given an immediate negative answer, by this theorem, whereas the extended Plücker relations would be cumbersome to handle or would yield no conclusive answer. A simple example shows this: A curve of order 13 can have 22 ordinary triple points; when these become origins of a single branch each, the Plücker relations reduce this number to 19. Now by the theorem just proved, curves of order $m-3-\mu$ with these points as double points must form a regular system for $k(k+1)\mu < m$; but the system of curves of order 9 with 19 double points has virtual dimension less than -1, hence 19 such points are impossible for an irreducible curve of order 13. As an example of the second type of restriction, consider curves of order 10, with points O each the origin of a single branch of order 3 and class 1; cubics simply through these points must form a regular system. There cannot be eleven such points, then; there are at most ten and if there are nine, they cannot be base points of a pencil of cubics.

13. *One branch of class greater than* 1. The two cases treated in detail make clear the fact that if we know the inequalities governing exponents of terms occurring in the equation of the adjoint surface ϕ_ν, we can describe the systems of curves $|C^d|$ cut out in the plane π, and enunciate for these systems theorems on regularity for determined orders. When O is the origin of a branch of order k and class c prime to k, the governing inequality * obtained by a transformation of the integral is of the form

$$(22) \qquad kn(\alpha+\beta-k+2-c)+nc(\beta+1)+k(k+c)(\gamma+1)>0.$$

Under this condition the curves $|C^d|$ must have the following properties: The generic C^0 has points O as $(k-1)$-fold points; all tangents at the point are arbitrary if the class $c \leq k/k-1$; one tangent coincides with the tangent to f if $k/k-1 < c \leq k/k-2$, etc. Further, the contact may be more than ordinary for any branch of C^0 which has contact with f, when c and k satisfy relations readily obtained from (22). All systems given by

$$d=0, 1, \cdots, [nc/k(k+c)]-1$$

have the points $O((k-1)$-fold; the next set, given by

* This inequality can be obtained directly from the description of the behavior of the adjoint surfaces at O as given by the Newton polyhedron for F, as well as by a transformation exactly analogous to that of paragraph 8: See W. V. D. Hodge, "The isolated singularities of an algebraic surface," *Proceedings of the London Mathematical Society*, ser. 2, Vol. 30, part 2 (1929).

$$d = [nc/k(k+c)], \cdots, [n(k+3c)/k(k+c)] - 1$$

have the points $O(k-2)$-fold, and so on. The systems $|C^0|, \cdots, |C^{n-2}|$ whose superabundances may be involved in expressing the irregularity of F do not necessarily include systems with points O as ordinary $(k-i)$-fold points with arbitrary tangents. The sequence contains a group of systems $|C^{\delta_1}|$ to $|C^{\delta}|$ having the points O as $(k-i)$-fold points and satisfying certain other conditions as to number of tangents coinciding with the tangent to f and as to the nature of the contact; these additional conditions decrease as we progress from δ_1 to δ, but they do not necessarily disappear entirely before the appearance of $|C^{\delta+1}|$ with $O(k-i-1)$-fold. However, if we denote by s_δ the superabundance of $|C^{\delta}|$, then a system $|C|$ of the same order having the points $O(k-i)$-fold with arbitrary tangents will have superabundance $s \leq s_\delta$, hence if $|C^{\delta}|$ must be regular, certainly $|C|$ is regular. In this way we are able to enunciate theorems analogous to those of part II on the regularity of curves of determined order with behavior at the points O specified.

III. Application to f having Points of different Multiplicities.

14. In I and II, the behavior of the curves $|C^d|$ is described at points O for any value of d. The theorems there are stated for the case of f with multiple points all of the same order, but when different multiplicities are present, it is necessary only to describe the behavior for separate multiplicities and then state the general theorem as the intersection of these theorems. Let $f(x, y) = 0$ be irreducible of order m, with ordinary multiple points $O_1, \cdots, O_r$ of orders $k_1, \cdots, k_r$ respectively. From paragraph 6 we deduce the theorem: *All complete systems of curves of order $m-3-\mu$ with the points O_i as $(k_i - 2 - \rho_i)$-fold points are regular provided*

$$k_i\mu < (\rho_i + 1)m. \qquad (i = 0, \cdots, r) \quad (\rho_i = 0, 1, \cdots, k_i - 2).$$

Suppose f has, in addition, points O_h each the origin of a single branch of order k_h and class 1. Each order k_i determines a set of critical values of d

$$d = [n/k_i], \quad [2n/k_i], \cdots, \quad [(k_i - 2)n/k_i], \qquad (i = 0, 1, \cdots, r)$$

for which the multiplicity of points O_i on C^d decreases by one; similarly, each k_h determines a set of values

$$d = [n/k_h(k_h+1)], \cdots, \quad [(k_h^2 - k_h - 1)n/k_h(k_h+1)]$$

for which the multiplicity of points O_h on C^d decreases by one. If these values are arranged in order of magnitude, the sequence $|C^0|, |C^1|, \cdots$ may be

described successively, for each time that a critical value of d is reached, the subscript will show which points change their behavior on $| C^d |$ for that value of d. The systems with base points thus specified will be regular for order $m-3-\mu$, where μ can be determined by the method of paragraph 6. For example, suppose f has cusps and ordinary triple points. The critical values of d arranged in order of magnitude give the behavior of $| C^d |$ as follows:

d	0	$[n/6]'$	$[n/3]$
Multiplicity of C^d at triple points	1	1	0
Multiplicity at cusps	1	0	0

Consequently, curves of order $m-3-\mu$ passing simply through the triple points and cusps are regular for $6\mu < m$; curves of order $m-3-\mu$ passing simply through the triple points only are regular for $3\mu < m$.

The nature of the curves $| C^d |$ occurring in the sequence for any given f depends on the relation of the multiplicities $k_1 \cdots k_r$ among themselves, since n may be taken as large as we please. If, for example, f has multiple points of orders k_1, k_2 respectively, where $k_1 = \rho k_2$, the multiplicity of $| C^d |$ at points O_1 will have reached $k_1 - 2 - \rho$ when the multiplicity at points O_2 decreases to $k_2 - 3$. The systems of curves obtained by this method are essentially different, then, from the curves obtainable by the process of successive adjunction to f.

IV. Curves with Non-cyclic Fundamental Groups.

15. *Sufficient conditions that the fundamental group of f be non-cyclic.* In his paper on cyclic multiple planes,* Zariski proved that if the fundamental group of a plane irreducible algebraic curve $f(x, y) = 0$, in general position with respect to the line at infinity, is cyclic, the surface $z^n = f(x, y)$ is regular for every positive n. From this and the expression of the irregularity in terms of the superabundances of the systems of curves through the cusps, he deduces a connection between the structure of the fundamental group and the position of the cusps of f when f has nodes and cusps only. In any case, a condition under which $z^n = f(x, y)$ is irregular will be a sufficient condition that f have a non-cyclic fundamental group.

When f has only ordinary k-fold points, the systems of curves $| C^d |$ fall into consecutive sets each with the points $O(k-i)$-fold; if any of these systems are superabundant, F will be irregular. From paragraph 6, we see that if $m-3-j$ is the maximum order for which the points O as $(k-i)$-fold base points give rise to a superabundant system, then

* O. Zariski, "On the irregularity, etc.," *loc. cit.*

$$(23) \quad k\lambda_i = (h+1)\{kj_i - (i-1)m\} - k - k\,[j_i m_0/m] + (i-1)m_1 + \sigma.$$

If $\lambda_i \geq 0$ for a particular value of i, no non-vanishing superabundance will appear due to those $|C^d|$ with points $O(k-i)$-fold; when $\lambda_i \geq 0$ for all values of i from 2 to $k-1$, no non-vanishing superabundance can appear in (12), and F is regular. From (23) we obtain

$$kj_i \geq (i-1)m \qquad (i = 2, 3, \cdots, k-1).$$

If $kj_i > (i-1)m$ for a particular i, the corresponding $\lambda_i \geq 0$; if then $kj_i > (i-1)m$ for all the values of i given, F is regular for $n \geq m$. It is then of necessity regular for values $n < m$, for if F were irregular for a value $n = n_1$, it would be irregular for n any multiple of n_1. For F to be irregular, therefore, it is necessary that at least one of the quantities

$$m,\ 2m,\ 3m, \cdots,\ (k-2)m$$

be divisible by k.

Suppose that for a particular value of i between 2 and $k-1$, we have $kj_i = (i-1)m$. From (23) we see that if $(i-1)m_1$ is not divisible by k, $\lambda_i \geq 0$; we must have not only $(i-1)m$ but also $(i-1)m_1$, and hence $(i-1)n$, divisible by k. When this is true, $\sigma = 0$ and $\lambda_i < 0$; we may state the result: *If $f(x, y) = 0$ is an irreducible algebraic curve of order m, with only ordinary k-fold points, the necessary and sufficient condition that the surface $z^n = f(x, y)$ $(n \geq m)$ be irregular is that $(i-1)m$ and $(i-1)n$ be divisible by k for at least one value of $i = 2, 3, \cdots, k-1$, and that putting $kj_i = (i-1)m$, the system of curves of order $m-3-j_i$ with the k-fold points of f as $(k-i)$-fold points be superabundant.*

For f having points of the second kind discussed, of course, more values of d enter as critical. We may state a set of sufficient conditions: for example,—*A sufficient condition that f, irreducible of order m, with points O_1 ordinary k-fold and points O_2 each the origin of a single branch of order k and class 1, have a non-cyclic fundamental group is that m be a multiple of $k(k+1)$, and that writing $m = k(k+1)j$, the systems of curves of order $m-3-j$ with $O_1(k-2)$-fold and $O_2(k-1)$-fold be superabundant.* The condition that f with nodes and cusps only have a non-cyclic fundamental group, mentioned at the beginning of this paragraph, is a special case of the above theorem. The set of sufficient conditions concludes with: *A sufficient condition that f, irreducible of order m, with points O_1 ordinary multiple points of order k and points O_2 each the origin of a single branch of order k and class 1, have a non-cyclic fundamental group is that $(k^2-k-1)m$ be divisible by $k(k+1)$, and that for $(k^2-k-1)m = k(k+1)j$, the curves of order $m-3-j$ passing simply through the points O_2 form a superabundant system.*

16. Suppose that f has ordinary multiple points $O_1, \cdots, O_r$ of orders $k_1, \cdots, k_r$ respectively (arranged in order of magnitude). The systems $|C^d|$ have the points O_i as $(k_i - 2 - \rho_i{}^d)$-fold base points and the irregularity of $z^n = f(x, y)$ is expressed in terms of the superabundances of these systems i. e.

$$q = s_0 + s_1 + \cdots + s_t \qquad t = [(k_1 - 2)n/k_1] - 1. \tag{24}$$

If F is to be irregular, it is necessary and sufficient that a superabundant system appear in this sequence. The actual set of values $\rho_i{}^d$ need not be calculated; we shall set up systems to compare with $|C^d|$ as follows: Let $\rho_1, \cdots, \rho_r$ be a set of r numbers, each associated with a k_i, where each ρ_i may have any value from 0 to $k_i - 2$. From (4) we know that $|C^d|$ will have O_i as $(k_i - 2 - \rho_i)$-fold point (at least), provided $d < [(\rho_i + 1)n/k_i]$. Evaluate this quantity for each pair ρ_i, k_i and denote by R, K that particular pair giving the minimum value δ for d;

$$[(R + 1)n/K] = \delta.$$

(If two or more pairs give the same minimum value for d, use any one of them as R, K.) Then for $|C^0|, \cdots, |C^{\delta-1}|$ the points O_i are of order $k_i - 2 - \rho_i{}^d \geq k_i - 2 - \rho_i$ for $i = 1, 2, \cdots, r$.

Now suppose that $m - 3 - j$ is the maximum order for which the systems of curves $|C|$ with the points O_i as $(k_i - 2 - \rho_i)$-fold base points are superabundant; systems of curves of the same order with O_i $(k_i - 2 - \rho_i{}^d)$-fold will then certainly be superabundant. Curves conditioned in this way occur in the first δ terms of the sequence $|C^d|$, so if the order $m - 3 - j$ is such that the number of terms preceding it is less than δ, a non-vanishing superabundance will occur in (24) and F will be irregular. The first system of order $m - 3 - j$ is preceded by $j(h + 1) - 1 - [jm_1/m]$ terms, where $n - m = hm - m_1$. Denoting the difference between this and δ by λ, we have

$$\lambda = j(h + 1) - 1 - [jm_1/m] - [(R + 1)n/K].$$

If $\lambda < 0$, F is irregular for there is a system $|C^d|$ of order $m - 3 - j$ given * by a value of $d < [(R + 1)n/K]$; it has the points O_i to multiplicity $\geq k_i - 2 - \rho_i$ and is therefore superabundant. Writing

$$\begin{aligned} (R + 1)n &= \tau K + \sigma, & (0 \leq \sigma < K), \\ jm_1 &= h_1 m + m_2, & (0 \leq m_2 < m), \end{aligned} \tag{25}$$

we obtain

$$K\lambda = (h + 1)\{Kj - (R + 1)m_1\} - K - Kh_1 + (R + 1)m_1 + \sigma. \tag{26}$$

* A system $|C^d|$ of order $m - 3 - j$ does occur for every j from 0 to $m - 1$, since by par. 2, there are $h + \epsilon_\mu$ systems of each order, and $h > 0$.

As before, $Kj \geq (R+1)m$ since λ cannot be less than zero for all large values of n, F being regular when n is a power of a prime.

Suppose $Kj > (R+1)m$ for a given set of values $\rho_1, \cdots, \rho_r$ defining R, K. If $h_1 = 0$, $K\lambda > -K$ and $\lambda \geq 0$. If h_1 is not zero, from (26) we have

$$Kj - (R+1)m = \{Kh_1 - (R+1)m\}j/h_1 + (R+1)m_2/h_1.$$

Since $m_1 < m$, $j > h_1$; hence even if $h = \sigma = 0$, $K\lambda > -K$ and again $\lambda \geq 0$. If then $Kj > (R+1)m$ for all possible sets of numbers $\rho_1, \cdots, \rho_r$, all the corresponding λ values are greater than or equal to zero. Under these conditions F may be shown to be regular as follows: If for one set $\rho_1, \cdots, \rho_r$ the corresponding $\lambda \geq 0$, we know that no non-vanishing superabundance appears due to curves with the points O_i exactly $(k_i - 2 - \rho_i)$-fold. For an arbitrary choice of $\rho_1, \cdots, \rho_r$, these curves are not necessarily curves $|C^d|$, and in this case, $\lambda \geq 0$ does not give information about the regularity of F. However, any system $|C^d|$ is described by some set $\rho_1, \cdots, \rho_r$, and if the λ defined for that set is ≥ 0, the system is not superabundant. We conclude that *if for every possible choice* $\rho_1, \cdots, \rho_r$ *the* λ *defined by* (26) *is* ≥ 0, *no* $|C^d|$ *is superabundant and* F *is regular.* For F to be irregular, then, it is necessary as well as sufficient that for at least one set of numbers $\rho_1, \cdots, \rho_r$, the λ defined be less than zero.

We saw that for $Kj > (R+1)m$, $\lambda \geq 0$; we must have $Kj = (R+1)m$ for some set $\rho_1, \cdots, \rho_r$. The λ for that set is given by

$$K\lambda = -K - K[(R+1)m_1/K] + (R+1)m_1 + \sigma.$$

It is clear that λ will be ≥ 0 unless $(R+1)m_1$ is divisible by K. In this case, $(R+1)n$ is also divisible by K, $\sigma = 0$, $\lambda < 0$, and F is irregular. We may state the result as follows: Given $f(x, y) = 0$, algebraic, and irreducible of order m, with $O_1, \cdots, O_r$ ordinary multiple points of orders $k_1, \cdots, k_r$ respectively. Let $\rho_1, \cdots, \rho_r$ be any set of r numbers, each ρ_i associated with a k_i and allowed to assume any value from 0 to $k_i - 2$, and denote by R, K a pair ρ_i, k_i which gives to $[(\rho_i + 1)m/k_i]$ its minimum value for $i = 1, 2, \cdots, r$. Then *a necessary and sufficient condition that* $z^n = f(x, y)$ $(n \geq m)$ *be irregular is that for at least one set of numbers* $\rho_1, \cdots, \rho_r$, *both* $(R+1)m$ *and* $(R+1)n$ *be divisible by* K, *and that for* j *defined by* $Kj = (R+1)m$, *the system of curves of order* $m - 3 - j$ *with the points* O_i *as* $(k_i - 2 - \rho_i)$*-fold base points be superabundant.* Since F is regular if f has a cyclic fundamental group, the condition given is a sufficient condition that the fundamental group of f be non-cyclic.

THE JOHNS HOPKINS UNIVERSITY.
BRYN MAWR COLLEGE.

SECOND NOTE ON THE CELESTIAL SPHERE.

By FRANK MORLEY.

1. *The caustic of a correspondence.* In a note on the Celestial Sphere [this Journal, Vol. 54 (1932), p. 276], I attempted to show the convenience of regarding the space around one as bounded by a sphere Ω. The points of Ω are named by numbers $x, y \cdots$. An arc orthogonal to Ω (or if preferred a line of the hyperbolic space) is named by its end points x and y.

To call Ω the Cayley absolute would be to begin at the wrong end—to pass from projective geometry to inversive geometry. The aim is rather to infer the properties of the included space from the known nature of its boundary.

We suppose now a correspondence between the points x and y, and that this is differentiable, so that for $x + dx$ we have in general $y + dy$.

We have a double infinity or congruence of arcs x, y. And the problem is to determine their envelope. This is, in a Euclidean space (when Ω is a point), a part of Hamilton's theory of systems of rays.

It is convenient to regard the sphere Ω as a plane. An arc x, y is then a semi-circle, say above the plane. The arc x, y intersects a consecutive arc $x + dx$, $y + dy$ when the pairs of points are on a circle and are interlaced. The cross-ratio $dxdy/(x-y)^2$ is then positive. We have then, when $d\theta$ is real,

$$dxdy/(x-y)^2 = (d\theta)^2; \tag{1}$$

thus the condition of intersection is that this differential invariant under homographies is to be invariant also under antigraphies (that is not altered by writing $\bar{x}$ for x) and is further to be positive. Geometrically stated, the elements dx and dy make opposite angles with $x - y$.

To find the point of intersection, we first find the point z where the join of x and y meets the join of $x + dx$ and $y + dy$. Taking the circle on which they lie as the base-circle, and replacing x and y by turns t and τ we are to have

$$z + \bar{z}t\tau = t + \tau$$

and

$$\bar{z}(\tau dt + t d\tau) = dt + d\tau.$$

Let $dt/d\tau = \mu e^{i\alpha}$, where μ is positive (the magnification).

Then $\tau^2 dt/t^2 d\tau = \mu e^{i\alpha}$

$$dt/d\tau = \pm \mu t/\tau.$$

For interlaced pairs, dt/t and $d\tau/\tau$ have the same sign, so that, as we have taken μ positive,

$$dt/d\tau = \mu t/\tau.$$

We have then

$$\bar{z}(1+\mu) = 1/t + \mu/\tau$$
$$z(1+\mu) = t + \mu\tau,$$

and therefore in general

$$z = (x + \mu y)/(1+\mu), \tag{2}$$

where

$$\mu = |\, dx/dy \,|\,.$$

The point z is thus the internal center of similitude for corresponding small circles around x and y.

On the normal to Ω at this point z, we take a distance ζ given by

$$\zeta^2 = (z - x)(\bar{y} - \bar{z}) \tag{3}$$

to obtain the intersection of the consecutive arcs. The envelope of the arcs is then given by (1), (2), and (3). For convenience let it be called the *caustic* of the correspondence.

2. *The caustic of a homography.* As a case in point, consider the homography

$$x = \kappa y.$$

Then $dx = \kappa dy$. Then from (1)

$$\frac{\kappa dy^2}{(\kappa - 1)^2 y^2} = (d\theta)^2,$$

where $d\theta$ is real. Hence $y = c \exp \lambda\theta$, where $\lambda = \nu\kappa - 1/\nu\kappa$, c a constant

$$x = \kappa c \exp \lambda\theta$$

$$z = \frac{\kappa + \mu}{1 + \mu}\, c \exp \lambda\theta$$

and

$$\zeta^2 = \frac{\mu(1-\kappa)(1-\bar{\kappa})}{(1+\mu)^2}\, c\bar{c} \exp (\lambda + \bar{\lambda})\theta,$$

that is $\zeta^2 = c_1 z\bar{z}$, where c_1 is a constant. This is a right cone. The arcs

then all touch this cone; the direction of an arc at contact makes a constant angle with the generator.

For the general homography

$$\alpha xy + \beta x + \gamma y + \delta = 0,$$

it follows that the arcs touch an inverse of a cone, that is a Dupin cyclide. The double points of the cyclide are the fixed points of the homography.

3. *The caustic of an antigraphy.* As a second case, consider the antigraphy

$$\alpha x\bar{y} + \beta x + \gamma\bar{y} + \delta = 0,$$

When there are fixed points, say 0 and ∞, this is

$$x = \mu\bar{y}.$$

The equations (1), (2), (3) are unaltered from (1)

$$\frac{\mu dy d\bar{y}}{(x-y)^2} = (d\theta)^2.$$

Hence $x - y$ is a real.

Hence x, y, and z are reals. Thus in this case the only arcs which can intersect consecutive arcs are those erected on the axis of reals. We have

$$\zeta^2 = (1-\mu)^2 z^2/4\mu$$

where z *is* a real; and the arcs touch two half-lines in the vertical plane on the axis of reals, the lines being images as to Ω. For any antigraphy with fixed points, the envelope of arcs is accordingly two arcs which meet Ω at the fixed points.

When the antigraphy has interchanging points 0, ∞, it may be taken

$$x\bar{y} = t^2$$

where t is a given turn.

From (1) we have

$$(d\theta)^2 = -\frac{x}{\bar{y}}\,\frac{dy d\bar{y}}{(x-y)^2},$$

so that

$$\frac{x}{\bar{y}(x-y)^2} = \frac{\bar{x}}{y(\bar{x}-\bar{y})^2},$$

that is $xy/(x-y)^2$ is a real.

Hence x and therefore y must be taken on the base circle. The envelope of the arcs is then a circle parallel to Ω, lying on the base sphere.

For an antigraphy with interchanging points, the envelope of arcs is accordingly a circle, such that the spheres on it, which touch Ω, touch it at the interchanging points.

In general for an algebraic correspondence $f(x, y) = 0$, that is for a Riemann surface, we shall have a caustic *surface*. The equation (1) giving at the place (x, y), or its correspondent (y, x), two directions at right angles, gives a natural dividing of the surface into orthogonal curves.

But for a correspondence $f(x, \bar{y}) = 0$, we shall have a caustic *curve*. In particular this is the case when the equation is self-conjugate, that is, when it is the image-system of an algebraic curve.

4. *The euclidean case.* When the infinity Ω becomes a point, the Dupin cyclide becomes a right cylinder and the arcs are all lines which touch this cylinder at a given angle.

For antigraphies in this case, the circles which cut Ω, other than orthogonally, disappear. We are left with a circle of the euclidean space. The arcs are lines such that, if 0 be the center of the circle, 1 the radius, p any interior point, they are perpendicular to the stroke $0, p$ and make with the disc an angle whose cosine is p.

ON THE COMMON POINTS OF TWO PLANAR CUBICS AND OF TWO PLANAR CYCLIDES.

By Frank Morley and W. K. Morrill.

We consider a pencil of cubics in a plane,

$$C_0 + \lambda C_\infty = 0.$$

The nine base points $(1, 2, 3, \cdots, 9)$ of the pencil are said to be associated and any one is determined by the other eight (supposed given). The problem is: Given eight real points and a ruler to construct the associated ninth.

Any quartic Q on the base points may be written as

$$C_0 L_\infty - C_\infty L_0 = 0$$

where L_0, L_∞ are linear.

The cubic $C_0 + \lambda C_\infty = 0$ meets Q, outside of the base points, at three points on a line of the pencil $L_0 + \lambda L_\infty = 0$; and this pencil of lines meets the quartic at a point c, the coresidual point of the base points as to Q. In particular if the quartic is two conics, one, B, on 16789, and the other, B', on 2345, any cubic of the system meets B' at two points, B at one point; these three points are on a line which meets B again at a fixed coresidual point c. Now there exist an infinity of conics on 2345 and to each there corresponds a point c on 16789, and conversely. In particular to a pair of lines on 2345, say 23 and 45, corresponds a point c_1 which we may indicate by:

$$16789 \mid 23 \mid 45.$$

This point is determined by 16789 and 23. Any cubic on these seven points meets the line 23 again at x, the conic again at y; and the line xy meets the conic again at the point c_1. Therefore we may denote c_1 by:

$$16789 \mid 23.$$

To construct it we may take for the cubic on the seven points a conic 26789 and a line 13. This can be done in ten ways. The points x and y are then given by Pascal's theorem.

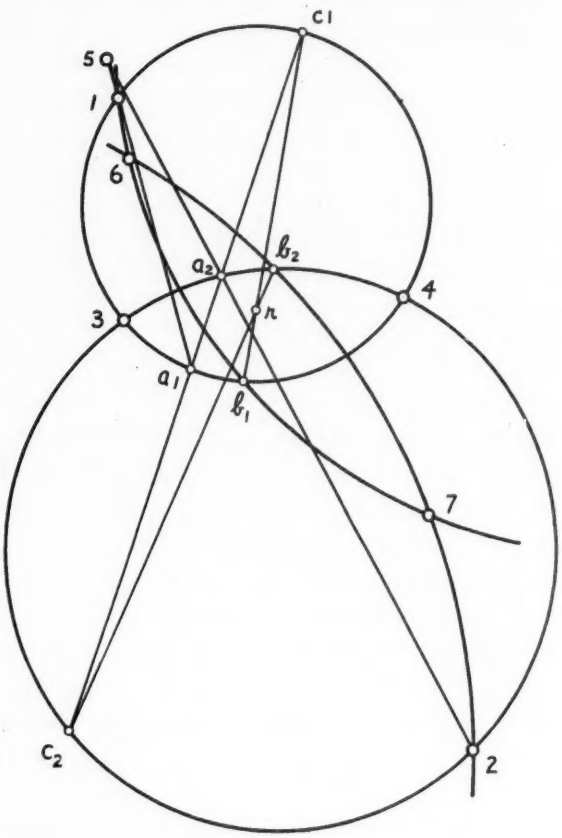
If 23 be the "circular points" and the conic 16789 be an ellipse, then using the eccentric angle θ the coresidual point is:

$$c_1 = \theta_1 + \theta_6 + \theta_7 + \theta_8 + \theta_9.$$

But the construction by Pascal's theorem fails.

The essential fact is that we have for the point c_1,

(1) $$16789 \mid 23 \equiv 16789 \mid 23 \mid 45 \equiv 16789 \mid 45.$$



Take now the two conics 16789 and 26789 and the point 5. We have the two degenerate cubics

$$16789 \times 25$$

and

$$26789 \times 15$$

meeting again at two points a_1 and a_2. The join of these points meets 16789 at c_1 or $16789 \mid 25$ and meets 26789 at c_2 or $26789 \mid 15$. Let us suppose now that 5 is the sought point. We recall that $16789 \mid 25$ is identically $16789 \mid 34$, and $26789 \mid 15$ is identically $26789 \mid 34$. Thus we have the construction of 5: Find the points $16789 \mid 34$ and $26789 \mid 34$. Their join

determines the points a_1 and a_2 on the conics, and $a_1 1$, $a_2 2$ meet at the required point 5.

There is the related inversive problem of the eight intersections of two cyclides (or bicircular quartics). Here seven points determine mutually an eighth. If 8 and 9 in the projective case be taken as the "circular points" the cubics are "circular cubics" or in the inversive plane cyclides on the point ∞. Thus the above construction becomes inversive if we replace 8 and 9 by the single point ∞. We have now the eight associated points 1234567∞ and wish to construct the point 5, when the other seven are given.

We take the pairs of circles

$$\begin{cases} 134 \text{ and } 234 \text{ on } 34. \\ 167 \text{ and } 267 \text{ on } 67 \end{cases}$$

Let the two radical axes 34 and 67 meet at r. Let 167 and 134 meet again at b_1. Let 267 and 234 meet again at b_2. Let rb_1 meet 134 again at c_1, and rb_2 meet 234 again at c_2. Let the line c_1c_2 meet 134 at a_1 and 234 at a_2. Then $a_1 1$ and $a_2 2$ meet at 5.

It is to be noticed that in the inversive problem we do not have to assume actual intersections. A pair of circles may or may not intersect. Thus 3 and 4 may be a common image pair. And it would not be difficult to handle the case when there are no actual intersections, that is when there are four common image-pairs.

For other solutions one may consult P. Serret, *Géométrie de Direction*, H. S. White, *Cubic Curves*, and the references given in H. F. Baker's *Principles of Geometry*.

THE RECTANGULAR FIVE-POINT.

By F. Morley and R. C. Yates.

I. An Involution-Form.

We take four points on a plane as the points μ_i, $1/\mu_i$ of a rectangular hyperbola. They are on a circle if

$$\mu_1\mu_2\mu_3\mu_4 = 1$$

and are orthocentric if

$$\mu_1\mu_2\mu_3\mu_4 = -1.$$

The Euler line of $\mu_1\mu_2\mu_3$ is then in rectangular coördinates

$$\begin{vmatrix} x & y & 1 \\ -1/s_3 & -s_3 & 1 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

and meets the curve at

$$\begin{vmatrix} \mu & 1/\mu & 1 \\ -1/s_3 & -s_3 & 1 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

that is, at $\mu + 1/s_3 = 0$, the orthocenter, and at a point μ_0 given by

$$\begin{vmatrix} \mu_0 & s_3 & 0 \\ 1 & s_3^2 & -s_3 \\ s_1 & s_2/s_3 & 3 \end{vmatrix} = 0$$

or, explicitly

$$(1)\quad \mu_0(\mu_1\mu_2 + \mu_2\mu_3 + \mu_3\mu_1 + 3\mu_1^2\mu_2^2\mu_3^2) = 3\mu_1\mu_2\mu_3 + (\mu_1 + \mu_2 + \mu_3)\mu_1^2\mu_2^2\mu_3^2.$$

Now this is for a given μ_0 an involution-form. That is, for given μ_1 and μ_2 we have three values of μ_3, say μ_3, μ_4, μ_5 which are symmetrically related to μ_1 and μ_2. For we have from (1)

$$\mu_1^2\mu_2^2(\mu_3 + \mu_4 + \mu_5) = 3\mu_0\mu_1^2\mu_2^2 - \mu_1^2\mu_2^2(\mu_1 + \mu_2)$$
$$\mu_1^2\mu_2^2(\mu_3\mu_4 + \mu_4\mu_5 + \mu_5\mu_3) = 3\mu_1\mu_2 - \mu_0(\mu_1 + \mu_2)$$
$$\mu_1^2\mu_2^2\mu_3\mu_4\mu_5 = \mu_0\mu_1\mu_2$$

that is, writing s_i for a product-sum of all five,

$$(2) \qquad s_5 = \mu_0$$

$$(3) \qquad s_1 = 3s_5$$

$$(3') \qquad s_4 = 3.$$

Five points on a rectangular hyperbola so related give ten triangles, as $\mu_1\mu_2\mu_3$; the Euler lines of the ten all meet at the point μ_0.

The equations (3) and (3′) are central in the paper of Mrs. Dean * so that we have here another way of approach to the theorems there given. It is there shown that the relations (3) and (3′) belong to the six points, μ_i and ∞. A six-point a_i is completed by drawing the six biquadratics which have a double point at one point and are on the other five. And in the case in question each curve cuts itself at right angles, or is rectangular. It is fair then to call this six-point *rectangular*, and to call a five-point, which with ∞ makes a rectangular six-point, also rectangular. Our five-point μ_i is then rectangular, in the sense that the 'circular cubics' on the five points and with a double point at one of them, will all cut themselves at right angles.

II. Neuberg's Invariant.

For four points, the condition that the four Euler lines meet at a point is found by eliminating μ_5 from (3) and (3′) to be

$$\sigma_1\sigma_3 - 10\sigma_4 + 3\sigma_4^2 + 3 = 0.$$

It is shown (p. 594 of the memoir cited) that this invariant is a factor of Neuberg's invariant for a four-point:

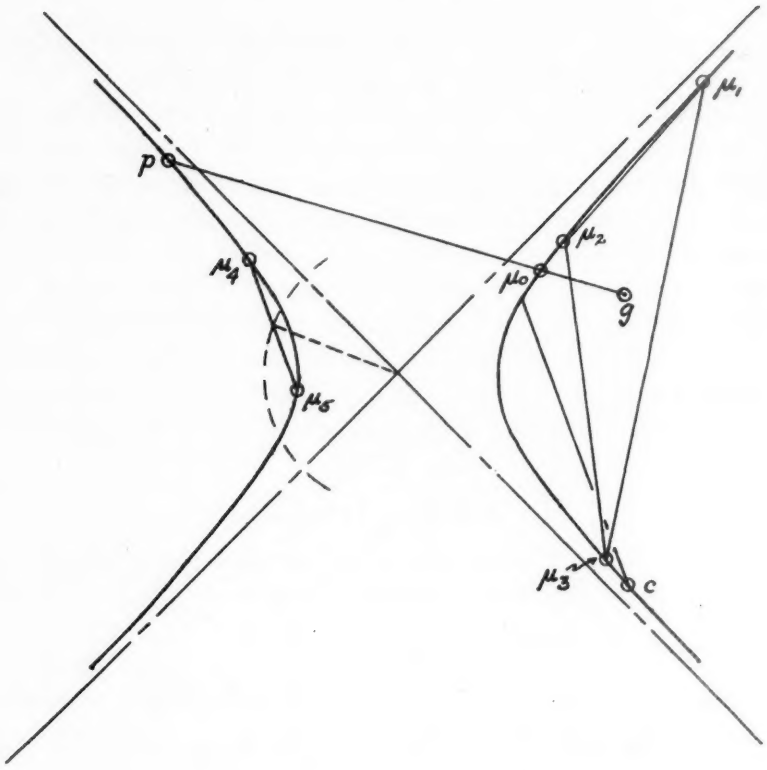
$$\begin{vmatrix} \lambda_{12}\lambda_{34} & \lambda_{12}+\lambda_{34} & 1 \\ \lambda_{23}\lambda_{14} & \lambda_{23}+\lambda_{14} & 1 \\ \lambda_{31}\lambda_{24} & \lambda_{31}+\lambda_{24} & 1 \end{vmatrix},$$

where λ_{ij} is the squared distance of two points. Accordingly, we may say that when the Euler lines all meet, for a four-point, the Neuberg invariant vanishes, or else the four-point is concyclic.

III. Construction of the Rectangular Five-Point.

It is of interest to construct a rectangular five-point. We take on a rectangular hyperbola three points μ_1, μ_2, μ_3 and mark the orthocenter p, the centroid g, and the point μ_0 where their join meets the curve again.

* *American Journal of Mathematics*, Vol. 52 (1930), p. 592.



We have from (2) and (3)

$$\mu_0 - (\mu_1 + \mu_2 + \mu_3)/3 = (\mu_4 + \mu_5)/3$$

and from (2) and (3′) the same in the reciprocals $1/\mu$. Hence, if we draw from the center of the hyperbola a stroke $3/2$ that from g to μ_0 we have the mid-point of μ_4 and μ_5. And from (2) the chord μ_4, μ_5 is parallel to the chord μ_0, c where c is the point $1/\mu_1\mu_2\mu_3$.

THE QUARTIC SPACE INVOLUTORIAL TRANSFORMATIONS WITH A DOUBLE CONIC.*

By L. A. Dye and F. R. Sharpe.

1. *Introduction.* Some of the birational transformations of space which are determined by quartic surfaces with a double conic, are briefly developed by Cremona's method in a paper by Aroldi.† Among these transformations there is a $(4-4)$ transformation with a double conic, a rational quartic curve meeting the conic in four points, and a single isolated fundamental point. Miss Aroldi gives a short synthetic discussion of the fundamental elements and of the principal surfaces of the transformation but does not derive the analytical form nor does she consider the possibility of the transformation being involutorial.

In this paper the equations of the involutorial quartic transformation with a double conic and a rational quartic curve are found by showing that the transformation is the product of a quadratic transformation and a bilinear cubic transformation. If the quadratic transformation is involutorial and certain restrictions are imposed on the cubic transformation, then the quartic transformation is involutorial and its equations are obtained by this method.

It is also shown that there is another type of involutorial transformation with a double conic which is the transform of a linear involutorial transformation by a quadratic transformation.

2. *The Cremona transformation.* $F_4 : C_2^2 + C_4(p=0) + 0$. Two quartic surfaces F_4 in a three space S_3 having a double conic C_2 in common, meet in a residual curve of order eight, C_8. By the formulas of Noether ‡ this C_8 meets C_2 in eight points and is of genus five. If the two surfaces have a rational quartic curve L_4 in common, meeting C_2 in four points, the C_8 consists of L_4 and a C_4 which, by Noether's formulas, meets C_2 and L_4 in four and six points respectively and is of genus zero. Conversely through C_4 passes ∞^4 quartic surfaces having C_2 as a double conic, such that any two meet in an L_4 which meets C_4 in six points and C_2 in four points. Three of

* Presented to the Society, December 28, 1931.

† G. Aroldi, *Giornale di Matematiche di Battaglini*, ser. 3, Vol. 58 (1920), pp. 175-192.

‡ M. Noether, *Annali di Matematica*, ser. 2, Vol. 5 (1871), pp. 163-177.

these quartic surfaces meet in two points not on C_2 or C_4, so that if one point is fixed, there is a homaloidal web of surfaces, $F_4 : C_2^2 + C_4 + 0$. There exists therefore a Cremona transformation, T_4 between two spaces S_3 and S'_3 by which a plane of S'_3 corresponds to a surface of the web of F_4 in S_3.

If the web of $F_4 : C_2^2 + C_4 + 0$ of the transformation T_4 is subjected to a quadratic involutorial transformation I_2 having C_2 and 0 as fundamental elements, a web of cubic surfaces $F_3 : C_2 + C'_4 (p = 0)$ is obtained. The composite curve $C_2 + C'_4$ is a $C_6 (p = 3)$, and the web of cubic surfaces define a bilinear Cremona transformation T_3. Hence the quartic transformation may be obtained as the product of a bilinear T_3 and a I_2 having a conic in common. The variable curve of intersection of two cubic surfaces of the web of F_3 is a C_3 meeting C'_4 in six points and C_2 in two points. The ∞^1 trisecants of C'_4 are P-lines ($P =$ principal) and they generate a quadric surface P_2 containing C'_4. The bisecants of C'_4 which meet C_2 are also P-lines, and their locus is a ruled surface of order six which contains C_2 as a triple conic and C'_4 as a double quartic, $P_6 : C_2^3 + C'^2_4$. These two P-surfaces make up the jacobian $J_8 : C_2^3 + C'^3_4$. Any cubic surface of the web meets P_2 in C'_4 and two P-lines, and meets P_6 in C_2^3, C'^2_4, and four P-lines. Hence P_2 corresponds to a conic C_2'', P_6 corresponds to a quartic curve C_4'', and the fundamental curve in S'_3 consists of the conic C_2'' and the rational quartic curve C_4'' which meets C_2'' in four points. The transformation T_3^{-1} is therefore similar to T_3.

Since $T_4 \equiv T_3 I_2$, it follows that if T_4 is involutorial $T_3 I_2 \equiv I_2 T_3^{-1}$, and hence that T_3^{-1} must have $C_2 + C_4$ as fundamental sextic, and that 0 must be invariant under T_3^{-1}. Amongst the web of F_3 there is a net of $F_3 : C_2 + C'_4 + 0$ which, by I_2, go into the plane of C_2 and a net of $F_3 : C_2 + C_4 + 0$. Amongst the web of F_2 there is a net of quadrics, each consisting of a plane through 0 and the plane of C_2, which are transformed by T_3^{-1} into a net of $F^3 : C_2 + C_4 + 0$ and the plane of C_2. Hence the net of planes through 0 are transformed by T_4 into the net of cubics $F_3 : C_2 + C_4 + 0$ and the plane of the conic. There is a composite cubic surface in the web of F_3 which consists of the plane of the conic and the unique quadric surface $P_2 : C'_4$. Under I_2 this goes into a unique quartic surface $P_4 : C_2^2 + C_4 + 0^2$. The plane of C_2 is transformed by I_2 into the cone $K_2 : C_2 + 0^2$ which goes by T_3^{-1} into the quartic surface $P_4 : C_2^2 + C_4 + 0^2$. Therefore by the involution T_4 the plane of the conic must go into this quartic surface.

3. *The web of* $F_3 : C_2 + C'_4$. The transformations T_3 and T_3^{-1} can be derived from three equations bilinear in (x) and (x'),

$$(1)\qquad \begin{aligned} u_1x'_1 + u_2x'_2 + u_3x'_3 + u_4x'_4 &= 0,\\ v_1x'_1 + v_2x'_2 + v_3x'_3 + v_4x'_4 &= 0,\\ w_1x'_1 + w_2x'_2 + w_3x'_3 + w_4x'_4 &= 0. \end{aligned}$$

When we put $x_4 = 0$, the four cubic curves, obtained by equating to zero the four third order determinants of the matrix

$$(2)\qquad \left\| \begin{matrix} u_1 & u_2 & u_3 & u_4 \\ v_1 & v_2 & v_3 & v_4 \\ w_1 & w_2 & w_3 & w_4 \end{matrix} \right\|,$$

must be of the form $(a_1x_1 + a_2x_2 + a_3x_3)K_2 = 0$, where $K_2 = 0$ is the cone $K_2 : C_2 + 0^2$. The two surfaces

$$\begin{aligned} u_1(v_3w_4 - v_4w_3) + v_1(u_4w_3 - u_3w_4) + w_1(u_3v_4 - u_4v_3) &= 0,\\ u_2(v_3w_4 - v_4w_3) + v_2(u_4w_3 - u_3w_4) + w_2(u_3v_4 - u_4v_3) &= 0, \end{aligned}$$

obtained from (1), determine a pencil of $F_3 : C_2 + C'_4$, passing through a residual C_3 which meets the conic $C_2 \equiv x_4 = K_2 = 0$ in two points P and Q and the plane $x_4 = 0$ in one other point R. If we put $x_4 = 0$ we get, by putting certain restrictions on u_1, u_2, v_1, v_2, w_1, and w_2, the conic C_2 through P and Q, and a pencil of lines through R. The matrix (2) is then of the required form. By linear combinations of rows and columns and by a linear transformation on (x), this matrix can be reduced to a normal form,

$$(3)\qquad \left\| \begin{matrix} 0 & -x_3 & x_2 & c_1x_1 \\ x_3 & 0 & -x_1 & c_2x_2 \\ -x_2 & x_1 & 0 & c_3x_3 \end{matrix} \right\|.$$

The four cubic equations obtained from it are $x_1K_2 = 0$, $x_2K_2 = 0$, $x_3K_2 = 0$, and a fourth which is identically equal to zero. The cone K_2 has the form $c_1x_1^2 + c_2x_2^2 + c_3x_3^2 = 0$. From (3) a general form for (1) can be obtained in the following manner; replace the first three columns in (3) by the three columns of the determinant which is the product of

$$\left| \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right| \equiv D \quad \text{and} \quad \left| \begin{matrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{matrix} \right|,$$

and add terms in x_4 with coefficients from the elements of

$$\left| \begin{matrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{matrix} \right| \equiv \Delta.$$

The three bilinear equations now have the form

$$\begin{aligned}
&(a_{13}x_2 - a_{12}x_3 + b_{11}x_4)x'_1 + (a_{23}x_2 - a_{22}x_3 + b_{12}x_4)x'_2 \\
&\qquad + (a_{33}x_2 - a_{32}x_3 + b_{13}x_4)x'_3 + c_1x_1x'_4 = 0, \\
&(-a_{13}x_1 + a_{11}x_3 + b_{21}x_4)x'_1 + (-a_{23}x_1 + a_{21}x_3 + b_{22}x_4)x'_2 \\
&\qquad + (-a_{33}x_1 + a_{31}x_3 + b_{23}x_4)x'_3 + c_2x_2x'_4 = 0, \\
&(a_{12}x_1 - a_{11}x_2 + b_{31}x_4)x'_1 + (a_{22}x_1 - a_{21}x_2 + b_{32}x_4)x'_2 \\
&\qquad + (a_{32}x_1 - a_{31}x_2 + b_{33}x_4)x'_3 + c_3x_3x'_4 = 0.
\end{aligned}$$

The equations of T_3 as obtained from these bilinear equations are

$$\begin{aligned}
T_3 &: x'_i = (A_{ij}x_j)K_2 + M_ix_4 + (c_jB_{ji}x_j)x_4^2, \qquad (i = 1, 2, 3), \\
&\quad x'_4 = x_4[(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j) + Lx_4 + \Delta x_4^2],
\end{aligned}$$

where $K_2 \equiv c_1x_1^2 + c_2x_2^2 + c_3x_3^2$,

$M_i \equiv$ a quadratic ternary form in x_1, x_2, x_3,
$L \equiv$ a linear ternary form in x_1, x_2, x_3,
$(A_{ij}x_j) = A_{i1}x_1 + A_{i2}x_2 + A_{i3}x_3$,
$(c_jB_{ji}x_j) = c_1B_{1i}x_1 + c_2B_{2i}x_2 + c_3B_{3i}x_3$, etc.

The A_{ij}, B_{ji} are cofactors of a_{ij}, b_{ij} in D and Δ respectively. The equations of T_3^{-1} are

$$\begin{aligned}
T_3^{-1} &: x_i = \phi'_i = (a_{ji}x'_j)K'_2 + M'_ix'_4 + [c_1c_2c_3(b_{ij}x'_j)x'^2_4]/c_i, \ (i = 1, 2, 3), \\
&\quad x_4 = x'_4\phi'_4 = x'_4[c_1(a_{j1}x'_j)^2 + c_2(a_{j2}x_j)^2 + c_3(a_{j3}x_j)^2 + c_1c_2c_3x'^2_4],
\end{aligned}$$

where $K'_2 \equiv (b_{1j}x'_j)(a_{j1}x'_j) + (b_{2j}x'_j)(a_{j2}x'_j) + (b_{3j}x'_j)(a_{j3}x'_j)$,

$M'_i \equiv$ a quadratic ternary form in x'_1, x'_2, x'_3,
$(a_{ji}x'_j) = a_{1i}x'_1 + a_{2i}x'_2 + a_{3i}x'_3$,
$(a_{j1}x'_j)^2 = (a_{11}x'_1 + a_{21}x'_2 + a_{31}x'_3)^2$, etc.

Let the involutorial quadratic transformation I_2 be

$$\begin{aligned}
I_2 &: x'_i = x_ix_4, \qquad (i = 1, 2, 3), \\
&\quad x'_4 = K_2 = c_1x_1^2 + c_2x_2^2 + c_3x_3^2.
\end{aligned}$$

The quartic transformation obtained from the product T_3I_2 is

$$\begin{aligned}
T_4 &: x'_i = x_4[(c_jB_{ji}x_j)K_2 + M_ix_4 + (A_{ij}x_j)x_4^2], \qquad (i = 1, 2, 3), \\
&\quad x'_4 = \Delta K_2^2 + LK_2x_4 \\
&\qquad + [(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j)]\,x_4^2,
\end{aligned}$$

and the inverse transformation T_4^{-1} obtained from $I_2T_3^{-1}$ is

$$T_4^{-1}: x_i = x'_4[(a_{ji}x'_j)K'_2 + M'_i x'_4 + \{c_1c_2c_3(b_{ij}x'_j)x'^2_4\}/c_i], \quad (i=1,2,3),$$
$$x_4 = [c_1\phi'^2_1 + c_2\phi'^2_2 + c_3\phi'^2_3]/\phi'_4.$$

The transformations T_4 and T_4^{-1} must have the same form when T_4 is involutorial, hence

(4_1) $K_2 \equiv K'_2$,
(4_2) $a_{ji} = c_j B_{ji}$,
(4_3) $(c_1c_2c_3b_{ij})/c_i = A_{ij}$,
(4_4) $M_i \equiv M'_i$,
(4_5) $(c_1\phi'^2_1 + c_2\phi'^2_2 + c_3\phi'^2_3) \equiv [\Delta K_2^2 + LK_2x_4 + \{(b_{j1}x_j)(A_{1j}x_j) + (b_{j2}x_j)(A_{2j}x_j) + (b_{j3}x_j)(A_{3j}x_j)\}x_4^2]\phi'_4$.

If condition (4_2) is used to express the a_{ji} in terms of c_j and b_{ji}, then condition (4_3) requires that $\Delta = 1$. With these restrictions on a_{ij} and b_{ij}, condition (4_1) is satisfied if $b_{ij} = b_{ji}$, and no further restrictions are necessary in order that conditions (4_4) and (4_5) shall be satisfied. The quartic involutorial transformation now has the form

$$I_4: x'_1 = x_4[(c_jB_{j1}x_j)K_2 + x_4\{c_2(c_jB_{j2}x_j)(b_{j3}x_j) - c_3(c_jB_{j3}x_j)(b_{j2}x_j)\} + c_2c_3(b_{j1}x_j)x_4^2],$$
$$x'_2 = x_4[(c_jB_{j2}x_j)K_2 + x_4\{c_3(c_jB_{j3}x_j)(b_{j1}x_j) - c_1(c_jB_{j1}x_j)(b_{j3}x_j)\} + c_1c_3(b_{j2}x_j)x_4^2],$$
$$x'_3 = x_4[(c_jB_{j3}x_j)K_2 + x_4\{c_1(c_jB_{j1}x_j)(b_{j2}x_j) - c_2(c_jB_{j2}x_j)(b_{j1}x_j)\} + c_1c_2(b_{j3}x_j)x_4^2],$$
$$x'_4 = K_2^2 + x_4^2[c_2c_3(b_{j1}x_j)^2 + c_1c_3(b_{j2}x_j)^2 + c_1c_2(b_{j3}x_j)^2],$$

where B_{ij} are cofactors of b_{ij} in

$$\Delta \equiv \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} = 1, \qquad b_{ij} = b_{ji},$$

and $K_2 = c_1x_1^2 + c_2x_2^2 + c_3x_3^2$.

4. *The Cremona transformation* $F_4: C_2^2 + C_4(p=0) + 0_1^2 + 0_2$. If the quartic curve C_4 has a double point 0_1^2 it can be shown that the quartic surfaces F_4 which contain it and a double conic C_2, also have the double point 0_1^2. The web of quartic surfaces F_4 may be transformed by a quadratic transformation $T_2: C_2 + 0_1$ into a web of quadric surfaces $F_2: C'_2 + 0_2$, hence $T_4T_2 \equiv S_2$ or $T_4 \equiv S_2T_2^{-1}$; that is, the quartic transformation may be regarded as the product of two quadratic transformations. Miss Hudson *

* H. Hudson, *American Journal of Mathematics*, Vol. 35 (1913), pp. 183-188.

has discussed this case and has derived the equations of the transformation.

If T_4 is involutorial, then $S_2T_2^{-1} \equiv T_2S_2^{-1}$ and S_2^{-1} has the same fundamental elements as T_2^{-1}. If S_2 is

$$(5) \qquad \begin{aligned} x'_i &= (a_{ij}x_j)(a_{4j}x_j), && (i=1,2,3;\ j=1,2,3,4),\\ x'_4 &= H_2[(a_{1j}x_j),\ (a_{2j}x_j),\ (a_{3j}x_j)], \end{aligned}$$

then under the transformation S_2^{-1}

$$(6) \qquad \begin{aligned} (i_{ij}x_j) &\sim x'_ix'_4, && (i=1,2,3;\ j=1,2,3,4),\\ (a_{4j}x_j) &\sim K_2(x'_1, x'_2, x'_3). \end{aligned}$$

Since $S_2S_2^{-1} =$ Identity, we can substitute equations (6) in (5) and obtain

$$\begin{aligned} x'_i &= x'_ix'_4K_2(x'_1, x'_2, x'_3), && (i=1,2,3),\\ x'_4 &= H_2(x'_1, x'_2, x'_3)x'^2_4, \end{aligned}$$

hence $H_2 \equiv K_2$.

The transformation S_2^{-1} is therefore

$$S_2^{-1}:\ x_i = (A_{ji}x'_j)x'_4 + A_{4i}K_2, \qquad (i=1,2,3,4;\ j=1,2,3),$$

where the A_{ij} are cofactors of a_{ij} in the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14}\\ a_{21} & a_{22} & a_{23} & a_{24}\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \equiv A.$$

The transformation S_2^{-1} is evidently the product of a linear transformation T_1 of determinant A, and the quadratic transformation T_2,

$$\begin{aligned} T_2:\ x'_i &= x_ix_4, && (i=1,2,3),\\ x'_4 &= K_2(x_1, x_2, x_3). \end{aligned}$$

Hence if T_4 is involutorial, $T_2T_1T_2^{-1} \equiv T_2T_1^{-1}T_2^{-1}$, and T_1 must be involutorial. The transform of a linear involutorial transformation by a quadratic transformation is a quartic involutorial transformation with a double conic and a quartic curve with a double point. It has the form

$$\begin{aligned} I_4:\ x'_i &= [(a_{ij}x_j)x_4 + a_{i4}K_2][(a_{4j}x_j)x_4 + a_{44}K_2],\ (i=1,2,3;\ j=1,2,3,4),\\ x'_4 &= (c_1a_{14}^2 + c_2a_{24}^2 + c_3a_{34}^2)K_2 + 2[c_1a_{14}(a_{1j}x_j) + c_2a_{24}(a_{2j}x_j)\\ &\quad + c_3a_{34}(a_{3j}x_j)]x_4K_2 + [c_1(a_{1j}x_j)^2 + c_2(a_{2j}x_j)^2 + c_3(a_{3j}x_j)^2]x_4^2, \end{aligned}$$

where the a_{ij} are elements of $A = |\ a_{ij}\ |$, $a_{ij} = A_{ji}$, and

$$K_2 \equiv c_1x_1^2 + c_2x_2^2 + c_3x_3^2.$$

A RELATION BETWEEN METRIC AND EUCLIDEAN SPACES.*

By W. A. WILSON.

1. *Introduction.* A semi-metric space is a set of points such that with each pair of points there is associated a unique non-negative number which is called the distance between them and is zero if, and only if, the points are identical. As Menger points out in his various articles on metric geometry,† a metric space is merely a semi-metric space which has the property that any three points are congruent with some three points in a Euclidean plane.

It is a natural inference from the above, as well as from Menger's general theorems, that a metric space in which any four points are congruent with some four points of a Euclidean space more nearly approximates Euclidean space than does a simple metric space. This inference is further strengthened by certain other results obtained by Menger (See II, pp. 209, 213). The present article is devoted to a further discussion of this question; the principal results are given in §§ 11 and 12.

2. *Notation.* A Euclidean space of n dimensions will be denoted by E_n. If a set A in a metric space Z is congruent with a set A' in a metric space Z', we say that A *can be imbedded in* Z'. If for a particular value of n any n points of the metric space Z can be imbedded in some Euclidean space, we say that Z has the *n-point property.*‡ Obviously, if n points can be imbedded in any Euclidean space, they can be imbedded in E_{n-1}.

The term *congruent* has its usual meaning and the fact that A is congruent with A' is denoted by the notation $A \simeq A'$. If $A \subset B$, $A' \subset B'$, $A \simeq A'$, $B \simeq B'$, and in every case the correspondence of a point a of A with a point a' of A' (Notation: $a \sim a'$) in the congruence $A \simeq A'$ is preserved in the congruence $B \simeq B'$, we say that the first congruence is *contained in* or is a *sub-congruence of* the second. If in the preceding definition the

* Presented to the Society, March 26, 1932.

† K. Menger, I. "Untersuchungen über allgemeine Metrik," *Mathematische Annalen*, Vol. 100, pp. 75-163, and Vol. 103, pp. 466-501; II. "Bericht über metrische Geometrie," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, Vol. 40, pp. 201-219; III. "New Foundation of Euclidean Geometry," *American Journal of Mathematics*, Vol. 53 (1931), pp. 721-745. These will be referred to as I, II, or III. Many of the principles of Euclidean hyper-geometry used later are conveniently collected in III, pp. 727-729.

existence of the congruence $A \simeq A'$ implies the existence of one and only one congruence $B \simeq B'$ containing the given congruence, we say that the congruence $A \simeq A'$ *determines* the congruence $B \simeq B'$. If $\{a_i\}$ and $\{a'_i\}$ are two sets of points and we say that $\sum a_i \simeq \sum a'_i$, it will be understood that in this congruence $a_i \sim a'_i$ for each i. Similar remarks apply to congruences between two simple arcs ab and $a'b'$, two triangles abc and $a'b'c'$, two tetrahedra, and in general any two simplices.

A space Z is *complete* if any Cauchy sequence $\{a_i\}$ contained in Z converges to a unique point a in Z. If a and b are two points and c is a third point such that $ac + cb = ab$,* we say that c is *between* a and b. If for any two points of a metric space Z there is a third point between them, Z is called *convex*. If for any pair of points a and b of Z there are points c and d, such that b is between a and c and a is between b and d, Z is called *externally convex*.†

3. THEOREM. *In a metric space let $a_1, a_2, \cdots, a_n$ be n points which can be imbedded in E_{n-1}. If for some $k < n$ there are points $a'_1, a'_2, \cdots, a'_k$ in some E'_{n-1} such that $\sum_1^k a_i \simeq \sum_1^k a'_i$, then the n given points can be imbedded in E'_{n-1} so that the given congruence is preserved.*

Proof. By hypothesis there are n points $\{a_i''\}$ in E_{n-1} such that $\sum_1^n a_i \simeq \sum_1^n a_i''$. Also we have $\sum_1^k a_i'' \simeq \sum_1^k a'_i$. By a principle of Euclidean hypergeometry there is at least one congruence between E_{n-1} and E'_{n-1} which contains the congruence $\sum_1^k a_i'' \simeq \sum_1^k a'_i$. In the congruence $E_{n-1} \simeq E'_{n-1}$ let the images of the respective points $\{a_i''\}$, $k+1 \leqq i \leqq n$, be $\{a'_i\}$, $k+1 \leqq i \leqq n$. Then we have $\sum_1^n a'_i \simeq \sum_1^n a_i'' \simeq \sum_1^n a_i$, which was to be proved.

4. THEOREM. *In a metric space let $a_1, a_2, \cdots, a_n$ be n points which can be imbedded in E_{n-1}. If $k \leqq n$ and the set $\{a_i\}$, $1 \leqq i \leqq k$, can be imbedded in E_{r-2}, $r \leqq k$, then the set $\{a_i\}$, $1 \leqq i \leqq n$, can be imbedded in $E_{n-(k-r)-2}$.*

Proof. Let $\sum_1^n a_i \simeq \sum_1^n a'_i$, where the points $\{a'_i\}$ lie in an E_{n-1}. From

---

* Here, as often, ab denotes the distance between a and b. The use of ab also to denote a segment will cause no confusion.

† For general properties of convex spaces see Menger, I.

‡ This property is enjoyed for every value of n by any Euclidean space and by Hilbert space. See Menger, II, p. 215.

the second hypothesis we see that the set $\{a'_i\}$, $i \leqq k$, lies in an E_{r-2} contained in E_{n-1}. The space E_{r-2} and the points $\{a'_i\}$, $i > k$, which do not lie in E_{r-2} determine a Euclidean space of dimensionality at most $r - 2 + n - k$. Hence the theorem is proved.

Remark. Thus, if five points can be imbedded in E_4 and three of them can be imbedded in a line or four in a plane, all five can be imbedded in E_3; and if four of them can be imbedded in a line, all five can be imbedded in a plane.

5. *Properties of lines.* Let Z be a convex complete space which has the four-point property. If a and b are points of Z and $\{x\}$ is the set for which $ax + xb = ab$, it has been shown elsewhere that this set is congruent with a Euclidean segment of length ab. Similarly the set of points $\{x\}$ for which $ax + xb = ab$, or $ab + bx = ax$, or $xa + ab = xb$ is congruent to a Euclidean segment, ray, or open line and may be properly called a *line in* Z.

Other properties of lines in Z are: (a) any two points determine a line; (b) a line is determined by any two of its points; (c) three lines having one point in common can be imbedded in E_3; and (d) if a, b, and c are three points, the union of the segments ab, ac, and bc can be imbedded in E_2.

Finally, if Z is also externally convex, a line in Z is congruent with an open line in Euclidean space; i. e., it extends indefinitely in both directions.*

6. *Simplices in convex complete space.* Let Z be a convex complete space and $a_0, a_1, a_2, \cdots, a_k$ be $k+1$ points which can be imbedded in E_k but not in E_{k-1}. Let $a_1, a_2, \cdots, a_k$ be the vertices of a $(k-1)$-dimensional simplex S_{k-1} in Z. Let x be any point of S_{k-1}. Then the union of the segments $\{a_0x\}$ as x ranges over S_{k-1} is called a *k-dimensional simplex in* Z with vertices $a_0, a_1, a_2, \cdots, a_k$.

We must first show that this definition by recurrence has sense. We first note that by § 4 any r of the given points can be imbedded in E_{r-1} but not in E_{r-2}. A single point is called a 0-dimensional simplex and may be denoted by S_0. For $k = 1$ and $S_0 = a_1$, a 1-dimensional simplex is the union of all the segments $\{a_0a_1\}$, of which one at least exists, since Z is convex and complete. For $k = 2$ we can take the point a_0 and the one-dimensional simplex S_1 whose vertices are a_1 and a_2 and build at least one 2-dimensional simplex S_2. Continue in this way until we reach the simplex S_k.

---

* The definition of line is due to Menger and the above properties are readily deduced from Menger's theorems (See I, pp. 75-113) and the four-point property. For details see a paper by the author, "On Angles in Metric Spaces," which will be published in the *Bulletin of the American Mathematical Society*.

The next three sections are devoted to showing that in a convex complete space Z having the k-point property any n points $\{a_i\}$, $n \leqq k$, which cannot be imbedded in E_{n-2} have the following properties *provided that* $k \geqq 4$: (1) they are the vertices of a unique $(n-1)$-dimensional simplex S_{n-1} in Z; (2) if they are congruent with a set $\{a'_i\}$ in E_{n-1}, this congruence determines a unique congruence of S_{n-1} with the Euclidean n-dimensional simplex having the points $\{a'_i\}$ as vertices which contains the given congruence; (3) if Z is also externally convex, the n given points determine a unique sub-set H_{n-1} of Z such that $H_{n-1} \simeq E_{n-1}$ and this congruence is determined by the congruence in (2). We first note that the above statement is not true if $k = 3$. For, let Z be the surface of a sphere and a, b, and c be any three points not on the same great circle. If distances are geodetic distances on the spherical surface, Z is convex, complete, and metric. In general the three points are the vertices of a spherical triangle which conforms with the above definition of a two-dimensional simplex S_2, but it is not congruent with the plane triangle which has sides of lengths ab, ac, and bc. If two of the points are diametrically opposed, an application of the above definition gives absurd results.

For simplicity the words segment, triangle, and tetrahedron will be used for one-, two-, and three-dimensional simplices.

7. THEOREM. *Let Z be a convex complete space which has the four-point property. Let a, b, and c be three points which cannot be imbedded in E_1. Then Z contains exactly one triangle T whose vertices are a, b, and c, and, if a', b', and c' are points in any Euclidean space such that*

$$a + b + c \simeq a' + b' + c',$$

this congruence determines a congruence of T with the Euclidean triangle T', whose vertices are a', b', and c'.

Proof. By § 5 each pair of points are the ends of a unique segment. As the point x ranges over the segment bc let T be the union of the segments $\{ax\}$. Also, if x and y are distinct points of bc, $(ax)(ay) = a$. For otherwise a, x, and y would lie on a line by § 5 and, since $x + y \subset bc$, a would be on the line determined by b and c, contrary to hypothesis.

We now proceed to prove that the given congruence determines a congruence $T \simeq T'$. By § 5, $ab + bc + ac \simeq a'b' + b'c' + a'c'$. Let the points u and v of T lie on the segments ax and ay, respectively, where $x + y \subset bc$. In the congruence just given let $x \sim x'$ and $y \sim y'$. Then $ax = a'x'$ and $ay = a'y'$. On $a'x'$ there is just one point u' for which $au = a'u'$. Let $u \sim u'$. Likewise let $v \sim v'$. We must prove that $uv = u'v'$.

By § 5 we have $ab + by + ay \simeq a'b' + b'y' + a'y'$. For the same reason $ax + xy + ay \simeq a'x' + x'y' + a'y'$; then $av + vx + ax \simeq a'v' + v'x' + a'x'$; this gives $uv = u'v'$.* Hence $T \simeq T'$ and, by construction, this contains the given congruence. There is no other congruence $T \simeq T'$ containing the given congruence, for this would involve a congruence of T' with itself leaving the vertices invariant, but not itself an identical transformation.

This does not complete the proof. For we have yet to show that, if a triangle were built up with one of the other segments as a "base", it would be identical with T. To do this we prove that any set P of points of Z which contains a, b, and c and is congruent with T' in such a way that $a \simeq a'$, $b \simeq b'$, and $c \simeq c'$ is identical with T and that the congruence $P \simeq T'$ is the same as $T \simeq T'$. Let u' be any point of T', and $u \sim u'$ and $u'' \sim u'$ in the respective congruences $T \simeq T'$ and $P \simeq T'$. We must show that $u'' = u$.

Let $a'u'$ produced meet $b'c'$ in x'. In the above congruences let u, x and u'', x'' correspond to u', x', respectively. Now $bx = b'x'$, $cx = c'x'$, and $bx + cx = b'c' = bc$ by the first congruence; and $bx'' = b'x'$, $cx'' = c'x'$, and $bx'' + cx'' = b'c' = bc$ by the second. Hence $x'' = x$. In like manner $au = a'u'$, $ux = u'x'$, and $au'' + u''x = a'x' = ax$; whence $u'' = u$.

Thus there is exactly one triangle T with the vertices a, b, and c and one congruence $T \simeq T'$ determined by the congruence $a + b + c \simeq a' + b' + a'$.

8. *N-dimensional convex complete spaces.* Let Z be a convex complete space which has the four-point property; let $a_0, a_1, a_2, \cdots, a_n$ be $n + 1$ points which can be imbedded in E_n but not in E_{n-1}; let these points be the vertices of an n-dimensional simplex S in Z, which is unique and is congruent with an n-dimensional simplex S' in E_n with vertices $a'_0, a'_1, a'_2, \cdots, a'_n$ such that $\sum_0^n a_i \simeq \sum_0^n a'_i$; and let H be the set of points $\{x\}$ in Z such that each x is collinear with two points in S. Then H is called the *n-dimensional subspace of Z determined by* $\{a_i\}$; for brevity we call H an n-space.

If $n = 1$, an n-space is a line in Z. Properties of lines in Z were given in § 5; analogous results will now be obtained for n-spaces where $n > 1$.

THEOREM I. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points* $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. *If the set* $\{a_i\}$ *is congruent with a set* $\{a'_i\}$ *in* E_n, *the congruence* $\sum_0^n a_i \simeq \sum_0^n a'_i$ *determines a congruence of H with a sub-set of* E_n.

* For special positions of u and v, as on the same segment or on the parimeter, a similar, though, briefer, procedure gives the same result.

Proof. By the above definition there is an n-dimensional simplex S in Z whose vertices are $\{a_i\}$ and this is congruent with the n-dimensional simplex S' in E_n whose vertices are the points $\{a'_i\}$ in such a manner that $\sum_0^n a_i \simeq \sum_0^n a'_i$. If x is a point of H, let x be collinear with the points u and v of S and, to fix the ideas, let v lie between u and x. The congruence $S \simeq S'$ gives points u' and v' in S' corresponding to u and v. On $u'v'$ extended take the point x' so that v' is between u' and x' and $ux = u'x'$; then $vx = v'x'$. Let $H' = \{x'\}$. We now proceed to show that $H \simeq H'$ with $x \sim x'$.

Let w be any point of S different from u and v and let $w \sim w'$ by the congruence $S \simeq S'$. By the four-point property and § 3 there is a point x'' in E_n such that $u + v + w + x \simeq u' + v' + w' + x''$. Since $u'x'' = ux = u'x'$, $v'x'' = vx = v'x'$, and $u'x'' = u'v' + v'x''$, we have $x'' = x'$ and so $wx = w'x'$. Hence $S + x \simeq S' + x'$ for each x.

We now show that the correspondence of x' to x does not depend on the choice of u and v. Suppose that x were also collinear with r and s, where $r + s \subset S$; and, to fix the ideas, let x be between r and s. By the previous paragraph $rx = r'x'$ and $sx = s'x'$. As $rs = rx + sx$, this gives $r's' = r'x' + s'x'$; thus x' is on $r's'$ with $r'x' = rx$, just as it would have been chosen if we had started from the collinearity of r, s, and x.

Now let y be some other point of H collinear with points h and k of S and let k be between h and y. By the congruence $S + x \simeq S' + x'$ we know that $h + k + x \simeq h' + k' + x'$ and by the four-point property and § 3 there is a point y'' in E_n such that $h + k + x + y \simeq h' + k' + x' + y''$. Hence

$$h'y'' = hy = h'y', \quad k'y'' = ky = k'y', \quad \text{and} \quad h'y'' = h'k' + k'y''.$$

As $h'y' = h'k' + k'y'$, this gives $y'' = y'$. Then $x'y' = x'y'' = xy$. This gives both uniqueness of correspondence and congruence between H and H'.

There is no other congruence of H with H' containing the given congruence, for this would involve a congruence of H' with itself which would leave the $n + 1$ points $\{a'_i\}$ invariant but not be an identical transformation. This is impossible since the points $\{a'_i\}$ do not lie in an E_{n-1}.

THEOREM II. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. Then H is convex, complete, and closed, and contains every point collinear with any two of its points.*

Proof. We first show that, if x and y lie in H and z is collinear with x and y, then z lies in H. By the definition of H and Theorem I there is a set of points $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, in E_n, which are congruent with $\{a_i\}$ and

are the vertices of an n-dimensional simplex S' in E_n, and the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$ determines a congruence of S with S' and of H with a sub-set H' of E_n.

Let u' be any inner point of S and $u \sim u'$. Then by the application of the four-point principle as in the proof of Theorem I, we can easily show that, if z' is the point on the line through x' and y' such that $xz = x'z'$ and $yz = y'z'$, then $u + x + y + z \simeq u' + x' + y' + z'$, and so $uz = u'z'$. If v' is a point on $u'z'$ so near to u' that v' is within S', there is a corresponding point v in S and, as in the case of u, $vz = v'z'$. The relations $uz = u'z'$, $uv = u'v'$, and $vz = v'z'$ show that z is collinear with u and v and so belongs to H. This also proves that H is convex.

In order to prove the rest of the theorem we first note that, since Z is metric, every convergent sequence is a Cauchy sequence and, since Z is complete, every Cauchy sequence converges to a point in Z. Hence, in order to prove that H is closed and complete, we need only to show that, if $\{x_i\}$ is a sequence of points in H which converges to a point x, then x is in H.

Let u' be an inner point of the simplex S', $u \sim u'$, and $x_i \sim x'_i$. Then the sequence $\{x'_i\}$ converges to a point x' in E_n, $u'x'_i = ux_i$, and consequently $ux = u'x'$. Let r be so small that the n-dimensional sphere of radius r and center u' lies in S'. On each segment $u'x'_i$ take the point v'_i such that $u'v'_i = r$ and let $v_i \sim v'_i$. Then each v_i lies on ux_i and in S. Obviously the sequence $\{v'_i\}$ converges to a point v' and, if $v \sim v'$, $v_i \to v$. We then have $uv = u'v'$. Also $v'_ix'_i \to v'x'$, $v_ix_i \to vx$, and consequently $vx = v'x'$. The relations $uv = u'v'$, $vx = v'x'$, $ux = u'x'$, and $u'x' = u'v' + v'x'$ show that $ux = uv + vx$; hence u, v, and x are collinear. This proves that x lies in H.

THEOREM III. *Let Z be a convex complete space which has the four-point property and let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$. Let $\{b_i\}$, $i = 0, 1, 2, \cdots, n$, be points of H which determine an n-space K. Then $K = H$.*

Proof. By the definition of K the points $\{b_i\}$ are the vertices of an n-dimensional simplex S_n and any point x of K is collinear with two points of S_n. By Theorem II and the definition of a simplex $S_n \subset H$, since $\{b_i\} \subset H$. Applying Theorem II again, we see that x lies in H. Hence $K \subset H$ and in like manner $H \subset K$.

THEOREM IV. *Let the complete space Z be convex and externally convex, and have the four-point property. Let H be an n-space in Z determined by the points $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, which are congruent with the set $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, in E_n. Then this congruence determines a congruence of H with E_n.*

Proof. By Theorem I there is a sub-set H' of E_n such that $H \simeq H'$. We have to prove that $H' = E_n$. If x' is any point of E_n, there are points u' and v' of the simplex S' whose vertices are $\{a'_i\}$ such that u', v', and x' are collinear. To fix the ideas, let v' be between u' and x'. Since $u' + v' \subset H'$, there are corresponding points u and v in H. As Z is externally convex, the line through u and v contains a point x such that $ux = u'x'$ and v is between u and x. Hence x lies in H and corresponds to a point x'' in H'. Now $u'x'' = ux = u'x'$ and x'' lies on $u'v'$ produced. Then $x'' = x'$. That is, $E_n \subset H'$. As $H' \subset E_n$, the theorem is proved.

Many theorems analogous to those of Euclidean geometry can now be deduced from the four theorems just proved and will occur to the reader. Of these mention should perhaps be made of the fact that, if $k < n$ and $k+1$ points $\{a_i\}$ of an n-space H determine the k-space K, then $K \subset H$ and the congruence of H with E_n (or a sub-set H' of E_n) implies the congruence of K with the E_k (or a sub-set of E_k) determined by the points of E_n congruent with $\{a_i\}$. Also, if Z is convex and complete and has the four-point property, an n-space H_n determined by the set $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, contains every point x such that the set $x + \{a_i\}$ can be imbedded in E_n.

9. THEOREM. *Let Z be a convex complete space which has the four-point property. Let $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, be $n+1$ points which can be imbedded in E_n but not in E_{n-1}. Then Z contains exactly one n-dimensional simplex S whose vertices are the points $\{a_i\}$ and, if $\{a'_i\}$, $i = 0, 1, 2, \cdots, n$, are points in any Euclidean space such that $\sum_0^n a_i \simeq \sum_0^n a'_i$, this congruence determines a congruence of S with the n-dimensional Euclidean simplex S' whose vertices are $\{a'_i\}$.*

Proof. This has been proved in § 7 for $n = 2$. Hence the theorem is true by induction if its validity for $n-1$ implies its validity for n. Let us assume then that it is true for $n-1$, $n \geqq 3$, and every smaller integer.

Then any n of the points, as $a_1, a_2, \cdots, a_n$, are the vertices of a unique $(n-1)$-dimensional simplex in Z, which we denote by T_0, the subscript indicating here and later that all the points $\{a_i\}$ *except* a_0 are vertices. As the point x ranges over T_0, let S be the union of the segments $\{a_0x\}$, which are unique by § 5. Also, if x and y are distinct points of T_0, $(a_0x)(a_0y) = a$. For otherwise a_0, x, and y would be collinear by § 5 and a_0 would lie in the $(n-1)$-space in Z determined by the points $\{a_i\}$, $i = 1, 2, \cdots, n$. Such a space exists by § 8, Theorem I, in consequence of our assumption of the validity of the theorem now being proved for $n-1$. This, however, contradicts the hypothesis that $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, cannot be imbedded in E_{n-1}.

Now the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$ implies the congruences

$$(1^0) \qquad T_0 \simeq T'_0;$$
$$(1^1) \qquad T_1 \simeq T'_1;$$
$$(1^2) \qquad T_2 \simeq T'_2;$$
$$\cdots\cdots\cdots$$
$$(1^n) \qquad T_n \simeq T'_n;$$

since our theorem is supposed to be true for $n-1$ and every smaller integer. Here, of course, each T'_i denotes the Euclidean simplex with vertices corresponding to those of T_i. It should also be noted here that, since these congruences are determined by the congruence $\sum_0^n a_i \simeq \sum_0^n a'_i$, if T_{ij} is the $(n-2)$-dimensional simplex common to T_i and T_j, the congruence $T_{ij} \simeq T'_{ij}$ is a sub-congruence of both congruences (1^i) and (1^j).

Let x and y be any points of S; they lie on unique segments a_0u and a_0v, where u and v are in T_0 by the definition of S. In congruence (1^0) let $u \sim u'$ and $v \sim v'$, and let the line through u' and v' meet two of the $(n-2)$-dimensional faces of T'_0, say T'_{0i} and T'_{0j}, in points r' and s'; let $r \sim r'$ and $s \sim s'$. (A similar discussion holds for the special cases where u' and v' lie on the same $(n-2)$-dimensional face of T'_0.) By congruences (1^0), (1^i), and (1^j) we have $rs = r's'$, $a_0r = a'_0r'$, and $a_0s = a'_0s'$. Hence by § 7

$$(2) \qquad \text{triangle } (a_0, r, s) \simeq \text{triangle } (a'_0, r', s').$$

We already have a one to one correspondence between the segments $\{a_0u\}$ in S and the segments $\{a'_0u'\}$ in S'. By (2), $a_0u = a'_0u'$ and $a_0v = a'_0v'$. Hence there are points x' and y' on a'_0u' and a'_0v', respectively, for which $a_0x = a'_0x'$ and $a_0y = a'_0y'$. If we let $x \sim x'$ and $y \sim y'$, we have a one to one correspondence between S and S'. But by congruence (2), $xy = x'y'$. Hence we have proved that $S \simeq S'$. There is no other congruence $S \simeq S'$ containing the given congruence, for this would involve a congruence of S' with itself leaving the vertices invariant but not itself an identical transformation.

Now let P be any set in Z which contains the set $\{a_i\}$ and is congruent with S' in such a way that $a_i \sim a'_i$, $i = 0, 1, 2, \cdots, n$. Let x' be any point of S' and $x \sim x'$, $x'' \sim x'$ in the respective congruences $S \simeq S'$, $P \simeq S'$. If we can show that $x'' = x$, then $P = S$.

Let a'_0x' produced meet T'_0 in u'_1; $a'_1u'_1$ produced meet T'_{01} in u'_2; $a'_2u'_2$ produced meet T'_{012} in u'_3; etc., until we finally reach a point u'_{n-1} on the segment $a'_{n-1}a'_n$. (We may of course reach a point on an edge of S' earlier; if so, the following discussion is merely shortened.)

In the congruence $S \simeq S'$, let $x \sim x'$ and $u_i \sim u'_i$; in the congruence

$P \simeq S'$ let $x'' \sim x'$ and $u_i'' \sim u'_i$. Now $a_{n-1}u_{n-1} = a'_{n-1}u'_{n-1}$, $u_{n-1}a_n = u'_{n-1}a'_n$, and $a_{n-1}u_{n-1} + u_{n-1}a_n = a'_{n-1}a'_n$; while $a_{n-1}u''_{n-1} = a'_{n-1}u'_{n-1}$, $u''_{n-1}a_n = u'_{n-1}a'_n$, and $a_{n-1}u''_{n-1} + u''_{n-1}a_n = a'_{n-1}a'_n = a_{n-1}a_n$. Hence $u_{n-1} = u''_{n-1}$ and $a_{n-2}u_{n-1} = a_{n-2}u''_{n-1} = a'_{n-2}u'_{n-1}$. Since u'_{n-2} lies on $a'_{n-2}u'_{n-1}$, we can repeat the reasoning just employed and show that $u_{n-2} = u''_{n-2}$. Eventually we have $u_1 = u''_1$ and can then show that $x'' = x$, whence $P = S$.

Thus there is but one simplex with the vertices $\{a_i\}$, $i = 0, 1, 2, \cdots, n$; in other words the result is independent of the choice of "base" from the $(n-1)$-dimensional simplices T_0, T_1, $T_2, \cdots, T_n$.

This completes the proof. It is obvious that the above theorem is merely a generalization of Menger's theorem that any two points in a convex complete space are the ends of a segment. The four-point condition insures uniqueness, but as yet we need the assumption of an $(n+1)$-point condition to insure the existence of *any* simplex. We now turn to a discussion of this point.

10. Let $\{a_i\}$, $i = 0, 1, 2, \cdots, n$, be any $n+1$ points in a metric space and suppose that these points cannot be imbedded in E_{n-1}. Also, for any $k \leqq n$, suppose that any k of these points can be imbedded in E_{k-1}, but cannot be imbedded in E_{k-2}. It is shown by Menger's Theorems 7 and 8 (see I, pp. 132-133) that, if D_n is the determinant

$$\begin{vmatrix} 0 & 1 & 1 & . & 1 \\ 1 & 0 & (a_0a_1)^2 & . & (a_0a_n)^2 \\ 1 & (a_1a_0)^2 & 0 & . & (a_1a_n)^2 \\ . & . & . & . & . \\ 1 & (a_na_0)^2 & (a_na_1)^2 & . & 0 \end{vmatrix} ;$$

then the $n+1$ points can be imbedded in E_n if and only if sign $D_n \neq$ sign $(-1)^n$. (If $D_n = 0$, they can be imbedded in E_{n-1}.) If for any k of the points, $k \leqq n$, determinants $\{D_{k-1}\}$ similar to D_n are formed, no $D_{k-1} = 0$ and in each case sign $D_{k-1} \neq$ sign $(-1)^{k-1}$.

If the values of a_ia_j are fixed except for a_0a_1 and we set $x = a_0a_1$, the determinant when expanded has the form $Ax^4 + Bx^2 + C$, where A, B, and C are constants. Therefore a small change in the value of x does not alter the sign of D_n; also, if the change is small enough, none of the determinants $\{D_{k-1}\}$, $k \leqq n$, will be altered in sign. Hence, if $n+1$ points satisfy the conditions laid down at the beginning of this section, a change in the distance between any pair, if sufficiently small, will leave the validity of the conditions unaffected.

By expanding D_n we see that $A = -D_{n-2}$, where D_{n-2} is the determinant formed from the points $\{a_i\}$, $i \geqq 2$. By Menger's theorems* $A \neq 0$ and sign A

* Menger also gives an extensive discussion of the function $Ax^4 + Bx^2 + C$, which includes the following properties. See I, pp. 120-135, and II, pp. 737-743.

$= -$ sign $D_{n-2} =$ sign $(-1)^{n-2} =$ sign $(-1)^n$. If we allow $x = a_0a_1$ to vary, $D_n = Ax^4 + Bx^2 + C$ will vary and will have the sign of A for x large enough. Hence sign $D_n =$ sign $(-1)^n$ for x large enough. From the form of the function D_n we see that there are at most two non-negative real values α and β of x for which $D_n = 0$. In the next section it is shown that these always exist. Then, if $\alpha < \beta$, it is readily seen from the nature of the function $Ax^4 + Bx^2 + C$ that sign $D_n \neq$ sign $(-1)^n$ when $\alpha < x < \beta$.

Hence, if we can prove that the validity of the n-point condition implies that $\alpha < x < \beta$, it follows that the $n + 1$ points can be imbedded in E_n.

11. THEOREM. *Let Z be a convex complete space which is also externally convex and has the four-point property. Then Z has the n-point property for every integer n.*

Proof. It is sufficient to show that, if Z has the k-point property for every $k \leqq n$, where n is a fixed integer greater than or equal to 4, then Z has the $(n + 1)$-point property.

Let $a_0, a_1, a_2, \cdots, a_n$ be any $n + 1$ points of Z, $n \geqq 4$. If $k \leqq n$, any k of these points can be imbedded in E_{k-1}. If the other points are all collinear with a_0 and a_1, all can be imbedded in E_1. In the opposite case let them be numbered so that a_2 is not collinear with a_0 and a_1. Then by § 7, a_0, a_1, and a_2 are the vertices of a triangle S_2 and by § 8 they determine a plane H_2 in Z. If all the other points lie in H_2, all the points can be imbedded in E_2. In the opposite case let a_3 not be in H_2. Then no three of the points a_0, a_1, a_2, and a_3 are collinear by § 8, Theorem II, they are the vertices of a tetrahedron S_3 in Z by § 9, and they determine a 3-space H_3 in Z by § 8. If all the other points lie in H_3, all the points can be imbedded in E_3, and the theorem is proved because $n \geqq 4$. In the opposite case let a_4 not be in H_3. Then no four of the points a_0, a_1, a_2, a_3, and a_4 lie in the same plane; for, if a_4 were in the same plane as any three of the others, it would be in H_3, since H_3 contains the plane in Z determined by any three of its points. If $n = 4$, we then have the result that no n of the points can be imbedded in E_{n-2}, while any n can be imbedded in E_{n-1}.

If $n > 4$, the points a_0, a_1, a_2, a_3, and a_4 can be imbedded in E_4 by hypothesis, they are the vertices of a 4-dimensional simplex S_4, and they determine a 4-space H_4; etc. Eventually we reach a stage where all the points are imbedded in E_{n-1} or the points $a_0, a_1, a_2, \cdots, a_{n-1}$ determine a unique $(n-1)$-space H_{n-1} in Z which does not contain a_n. In the latter case no k of the points, $k \leqq n$, can be imbedded in E_{k-2}, while any n of them are the vertices of a unique $(n-1)$-dimensional simplex in Z.

Now $a_1, a_2, a_3, \cdots, a_n$ are the vertices of a unique $(n-1)$-dimensional simplex T_0 and $a_0, a_2, a_3, \cdots, a_n$ are the vertices of a unique $(n-1)$-dimensional simplex T_1. These have as a common face the $(n-2)$-dimensional simplex T_{01}.

Let E_{n-2} be an $(n-2)$-space in E_n containing points $a'_2, a'_3, \cdots, a'_n$ congruent with $a_2, a_3, \cdots, a_n$, and let T'_{01} be the Euclidean $(n-2)$-dimensional simplex congruent with T_{01} and having the vertices $a'_2, a'_3, \cdots, a'_n$. Since Z has the n-point property, it follows by § 3 that in any E_{n-1} containing E_{n-2} and contained in E_n there is a point a'_1 such that $\sum_1^n a_i \simeq \sum_1^n a'_i$ and $T_0 \simeq T'_0$, where T'_0 is the Euclidean $(n-1)$-dimensional simplex of vertices $\{a'_i\}$, $i = 1, 2, \cdots, n$. Likewise there is a point a'_0 such that

$$a_0 + \sum_2^n a_i \simeq a'_0 + \sum_2^n a'_i \quad \text{and} \quad T_1 \simeq T'_1.$$

Since a reflection about E_{n-2} is a congruence, each of the points a'_0 and a'_1 can be chosen in two ways. If a'_0 and a'_1 are on opposite sides of E_{n-2}, the segment $a'_0a'_1$ cuts E_{n-2} in a point b'. By § 8, Theorem IV, the $(n-1)$-space H_{n-1} determined by the vertices of T_0 is congruent with E_{n-1} by a congruence determined by that between the vertices of T_0 and T'_0. This congruence also determines the congruence $T_{01} \simeq T'_{01}$ and a congruence between the $(n-2)$-space H_{n-2} determined by the vertices of T_{01} and E_{n-2}. Let b be the point of H_{n-2} such that $b \sim b'$. Likewise the $(n-1)$-space K_{n-1} determined by the vertices of T_1 is congruent with E_{n-1}, $T_1 \simeq T'_1$, $T_{01} \simeq T'_{01}$, and $H_{n-2} \simeq E_{n-2}$. Hence in this congruence, too, $b \sim b'$. Since a'_0, b', and a'_1 are collinear, $a'_0a'_1 = a'_0b' + b'a'_1$. Also $a_0b = a'_0b'$ and $ba_1 = b'a'_1$ by the respective congruences $K_{n-1} \simeq E_{n-1}$ and $H_{n-1} \simeq E_{n-1}$. Hence the triangle axiom of metric geometry gives $a_0a_1 \leqq a_0b + ba_1 = a'_0a'_1$. Set $a'_0a'_1 = \beta$.

If a'_0 and a'_1 were so taken as to be on the same side of E_{n-2}, the discussion in the second paragraph of § 10 shows that there is no loss of generality in assuming that a'_0 and a'_1 are not equidistant from E_{n-2}. Then $a'_0a'_1$ produced will meet E_{n-2} in a point c', to which corresponds a point c in H_{n-2}, and in the same manner as in the previous paragraph we find that $a_0a_1 \geqq |\, a_0c - a_1c \,| = a'_0a'_1$. In this case set $a'_0a'_1 = \alpha$.

It is easy to show that $\alpha < \beta$. If $a_0a_1 = \alpha$ or $a_0a_1 = \beta$, the $n+1$ points $\{a_i\}$ can be imbedded in E_{n-1} and by Menger's Theorem 7 (see I, p. 132), the determinant D_n discussed in § 10 equals zero. We have seen in the two preceding paragraphs that $\alpha \leqq a_0a_1 \leqq \beta$. Hence by the results of § 10 the $(n+1)$-point condition is satisfied. This completes the proof.

12. THEOREM. *Let Z be a convex complete separable space which is also*

externally convex and has the four-point property. Then Z is congruent with some E_n or with Hilbert space.

Proof. In consequence of § 11 the space has the n-point property for every value of n. Hence the determinant D_n (see § 10) for any $n+1$ points has a sign different from that of $(-1)^n$ or is zero. If for every value of n there is some set of $n+1$ points for which $D_n \neq 0$, it then follows by a theorem of Menger * that Z is congruent to Hilbert space.

In the contrary event there is a smallest integer n such that, for every set of $n+2$ points, the determinant $D_{n+1} = 0$. Then every set of $n+2$ points can be imbedded in E_n (see I, p. 132). Consequently Z can be imbedded in E_n (see I, p. 128). But for some set of $n+1$ points the determinant $D_n \neq 0$ and so this set of points can be imbedded in E_n but not in E_{n-1}. Hence by § 8 they determine an n-space H_n congruent to E_n. As Z can be imbedded in E_n, $Z = H_n$ and the theorem is proved for this case.

13. Several questions are immediately suggested by the result just obtained. As remarked in § 1 a semi-metric space is metric if the 3-point condition is satisfied, *i. e.*, if the determinant D_2 formed from any three points is not positive. This happens to be the case if the triangle axiom is valid. In view of our last theorem it would be especially interesting to have an axiom similar to the triangle axiom and equivalent to the condition that the 4-point determinant $D_3 \geqq 0$.

The overwhelming effect of the 4-point condition lends new force to Menger's identification of metric space with semi-metric space satisfying the 3-point condition,—it shows how drastic the triangle axiom really is despite its simplicity and suggests the possibility that the proper point of departure in the study of abstract spaces is semi-metric rather than metric space.

It seems to the writer reasonable to suppose that the conditions given in the theorems of §§ 11 and 12 are stronger than needed. The results of this article originated in an attempt to find a convex space which satisfied the 4-point condition and was *not* Euclidean. It will be noted that, while the removal of the condition of external convexity in § 12 makes this theorem false, it seems not unlikely that there is still congruence with a convex part of some Euclidean or Hilbert space. The condition is of more serious importance in the demonstration of § 11, where it is needed to establish the existence of the point c in Z corresponding to c' in the next to the last paragraph. It may well be that the conclusion of the theorem can be established by some other method without using the condition of external convexity.

YALE UNIVERSITY.

* *Wiener Akademischer Anzeiger*, 1928, No. 12. See also III, p. 745.

ON THE CONSTRUCTION OF SIMPLE ARCS.

By GORDON T. WHYBURN.

1. In this paper it is proposed, in the first place, to give a simple, complete, and essentially new demonstration for the well-known theorem * that every compact locally connected continuum is arcwise connected. All previous proofs for this fundamental proposition with which the author is familiar are alike in that the arc is constructed by means of a monotone decreasing sequence of simple chains of open sets or regions. Due to the fact that two successive links in such chains necessarily overlap in more than one point, thereby leaving open the possibility that the next chain in the sequence may oscillate widely within a given chain, a certain element of elusiveness seems unavoidable when these chains are used to construct arcs; and in order to completely eliminate all such possibilities of oscillation and make the argument entirely convincing, some tedious details are necessary.

In the proof to be given below these difficulties are avoided by making use, in the construction, of chains of continua, called *regular chains*, defined as follows: C is a *simple regular chain* of continua joining two points a and b if C is the sum of a finite number of compact continua $a \subset X_1, X_2, \cdots, X_n \supset b$ such that any two successive links have exactly one common point, links that are not successive have nothing in common, and only the first and last links contain a and b respectively. If $\delta(X_i) < \epsilon$ for each i, we shall call C a simple regular ϵ-chain. We show first that any two points of a locally connected continuum are joinable by simple regular ϵ-chains of *locally connected continua* for any $\epsilon > 0$. Thus by simple recursion we can set up a monotone decreasing sequence of these chains from a to b such that the norm, ϵ, approaches 0, and their product is easily proved to be an arc from a to b. Clearly such regular chains represent better approximations to arcs than do the simple chains of open sets, because since each intersection point of two links in the chain separates the chain between a and b, every such point in any chain in the sequence must belong to the product set and thus to the final arc. No difficulties of oscillation can arise in this construction because, there being only one common point for any two successive links of a given chain, any later

* See R. L. Moore, *Transactions of the American Mathematical Society*, Vol. 17 (1916), p. 136; S. Mazurkiewicz, *Fundamenta Mathematicae*, Vol. 1 (1920), p. 201.

chain must pass through all of these intersection points in the same order as the links occur from a to b.

In the remaining sections of the paper we discuss other methods of constructing arcs and indicate ways of obtaining the more general arcwise connectivity theorem for complete locally connected spaces. In particular it is shown that the latter result may be obtained from a more general junction property of general locally connected spaces which can be demonstrated by means of regular chains.

2. We proceed now with our proof for the

ARCWISE CONNECTIVITY THEOREM. *Every two points a and b of a compact locally connected continuum M can be joined in M by a simple continuous arc.*

Proof. (1). For any $\epsilon > 0$, $M = \sum_1^n M_i$, where each M_i is a locally connected compact continuum of diameter $< \epsilon$.

For it is well known that M is the image under a uniformly continuous transformation T of the unit interval I. Thus if for the given ϵ, I is divided into a finite number of subintervals $I_1, I_2, \cdots, I_n$, each of diameter less than the corresponding δ_ϵ, then $T(I_1) = M_1$, $T(I_2) = M_2, \cdots, T(I_n) = M_n$ have the required properties.

(2). For any $\epsilon > 0$, any point x of M is an interior point (rel. M) of some compact locally connected subcontinuum $K_\epsilon(x)$ of M of diameter $< \epsilon$.

For, set $M = \sum_1^n M_i$ as in (1), where $\delta(M_i) < \epsilon/2$, and let $K_\epsilon(x)$ be the sum of all these continua M_i which contain x.

(3). Any two points a and b of a region * R in M can, for each $\epsilon > 0$, be joined in R by a simple regular ϵ-chain of compact locally connected continua.

For let S be the set of all points of R which can be so joined to a. Then S is open in R; for if x is any point of S, if C is a chain of the desired type in R from a to x, if X is the link of C containing x, and if $\sigma > 0$ is a number $< \epsilon - \delta(X)$ and also $< \rho[x, F(R) + C - X]$, then the chain obtained from C by replacing the link X by the set $X + K_\sigma(x)$ is an ϵ-chain of the desired type in R having x as an interior point. But S is also closed in R. For let y be any limit point of S in R. Let us choose $\sigma < \epsilon$ so that $K_\sigma(y) \subset R$. As y is an interior point of $K_\sigma(y)$ and thus $K_\sigma(y) \cdot S \neq 0$, hence there exists a

* By a *region* in M is meant any connected open subset of M.

simple regular ϵ-chain C of locally connected continua $a \subset X_1, X_2, \cdots, X_n \supset x$ in R from a to some point x of $K_\sigma(y)$ such that $X_i \cdot K_\sigma(y) = 0$ for $i < n$. Now if C contains y, then $y \subset S$. If not, then set $K_1 = K_\sigma(y)$, choose a number $\epsilon_1 < 1/2$ and also $< \rho(y, C)$ and, applying (1), express K_1 as the sum of a finite collection G_1 of locally connected ϵ_1-continua. Let H_1 be the sum of all the continua of G_1 which contain no point of X_n and let Y_1 be the component of H_1 containing y. Then G_1 contains at least one element, say K_2, which contains at least one point of Y_1 and at least one point of X_n. In general, for any $k > 1$, choose a positive number $\epsilon_k < 1/2^k$ and also $< \rho(Y_{k-1}, X_n)$ and express K_k as the sum of a finite collection G_k of locally connected ϵ_k-continua. Let H_k be the sum of all those continua of G_k which contain no point of X_n and let Y_k be the component of the set $Y_{k-1} + H_k$ which contains y. Then G_k contains at least one element, say K_{k+1}, which contains at least one point of each of the sets Y_k and X_n. Let $Y = \sum_1^\infty Y_k$. Then since $Y \cdot X_n = 0$, since each Y_k is closed, and $Y - Y_k \subset K_k$, it is seen at once that $\bar{Y} - Y = \prod_1^\infty K_k = p$, a single point; and thus $\bar{Y}$ is a compact locally connected ϵ-continuum having just the point p in common with X_n. Thus $[X_1, \cdots, X_n, \bar{Y}]$ is a simple regular ϵ-chain of compact locally connected continua in R from a to y. Hence every limit point y of S in R belongs to S and thus S is closed in R. As S is both open and closed in R and R is connected, we have $S = R$. Thus $b \subset S$ and (3) is proved.

We proceed now to construct an arc in M from a to b. By (3), a and b can be joined in M by a simple regular 1-chain C_1 of compact locally connected continua. Likewise for each $n > 1$, it follows by (3), since C_{n-1} is a locally connected continuum, that a and b can be joined in C_{n-1} by a simple regular $1/n$-chain of locally connected continua. Let $C = \prod_1^\infty C_n$. Then C is a simple arc from a to b, as will now be demonstrated. In the first place, since each C_n is a compact continuum containing $a + b$ and $C_n \supset C_{n+1}$, it follows that C is a compact continuum containing $a + b$. Furthermore, C is locally connected. To prove this, let x be any point of C and, for any n, let A_n be the sum of the (at most two) links of C_n which contain x and let B_n be the sum of those not containing x. Then since $A_n \cdot B_n$ consists of 0, 1, or 2 points according as A_n contains 2, 1, or 0 points of the set $a + b$, it follows that in any case $A_n \cdot C$ is connected; and since $\delta(A_n \cdot C) < 2/n$, it follows that C is locally connected. Finally, each point x of $C - (a + b)$ separates* a and

* That is, $C - x = C_a + C_b$, where C_a and C_b are mutually separated and contain a and b respectively.

b in C. For let P_n be the set of all points common to at least two links of C_n and let $P = \sum_1^\infty P_n$. Then clearly $P \subset C$ and every point of P separates a and b in C. Let x be any point of $C - P - a - b$. Then if x does not separate a and b in C, some component R of $C - x$ contains $a + b$. But by (3), R contains a compact continuum K containing $a + b$, and hence $K \supset P$; but this is impossible, because x is a limit point of P but not of K. Therefore x separates a and b in C and, by the Sierpinski definition, C is a simple arc from a to b.

3. *Remark.* Although in the above proof we showed that the set C is a simple arc from a to b by proving that it satisfies the Sierpinski definition of an arc, it is interesting to note that the method of construction of C lends itself readily to yield a direct proof that C is homeomorphic with the unit interval. For clearly the chain C_1 can, by choosing a sufficiently small norm and by combining the proper number of links, be so chosen that the set P_1 of all points common to at least two links of C_1 contains exactly $2^{v_1} - 1$ points, where v_1 is some integer. Similarly the chain C_2 may be so chosen that each of its links lies wholly in some link of C_1 and each link of C_1 contains exactly 2^{v_2} links of C_2, where v_2 is some integer uniform for all links of C_1, and thus so that the set P_2 of all points common to at least two links of C_2 contains exactly $2^{v_1+v_2} - 1$ points. We can continue in this way so that the set P_n of all points common to at least two links of C_n contains exactly $2^{v_1+v_2+\cdots+v_n} - 1$ points, and such that each link of C_{n-1} contains exactly 2^{v_n} links of C_n. Now if for each n we place the points of $a + P_n + b$ into correspondence with the numbers $m/2^{v_1+v_2+\cdots+v_n}$, $(m = 0, 1, 2, \cdots, 2^{v_1+v_2+\cdots+v_n})$, in the order in which they occur in C_n from a to b, we obtain a $(1-1)$ correspondence T between the set $\sum_1^\infty P_n$ and the set D of rational numbers of the form $m/2^n$. It is readily shown that both T and its inverse are uniformly continuous, and hence both may be extended to their limit points, thus giving a homeomorphism between C and the interval $(0, 1)$.

4. *Another method.* We note also that the arcwise connectivity theorem may be obtained in the following manner. First prove the theorem for the case of continua which are regular in the sense of Menger-Urysohn. This case results at once from the fact that any two points in a compact continuum can be joined in that continuum by an irreducible subcontinuum, because any subcontinuum of a regular curve is regular and hence locally connected, and any locally connected continuum irreducible between two points is an arc

joining these two points.* (The same argument proves the theorem for the case of hereditarily locally connected continua, *i. e.*, continua containing only locally connected subcontinua.) This case established, the general theorem is obtained by showing that any two points in a locally connected continuum lie together in a regular subcontinuum. This is done by constructing the set C as in § 2. That C is regular is self evident.

5. *Extension to complete spaces.* We turn now to the well-known † extension of the arcwise connectivity theorem to complete spaces. This is embodied in the following proposition which we shall call the

GENERALIZED ARCWISE CONNECTIVITY THEOREM. *Every two points a and b of a connected and locally connected complete space S can be joined in S by a simple continuous arc.*

In the opinion of the author this more general theorem can best be proved by reducing the general case back to the case of compact spaces M treated in § 2 by means of the following proposition: ‡

(A). *If K is any self-compact subset of a connected and locally connected metric space N, then there exists a subset H of N which is dense in K and which is the image under a uniformly continuous transformation of the set of all dyadic rational numbers on the unit interval.*

As has been pointed out by the author ‡ this theorem (A) yields at once

---

* A proof for this fact may be found in Hausdorff's *Mengenlehre* (1927), p. 222. It does not seem to be generally recognized that in this characterization of simple arcs no assumption whatever of compactness is necessary. In other words, *a metric space C is a simple arc from a to b if and only if* (β) *C is connected and contains a and b but no closed and connected proper subset of C contains both a and b and* (ϵ) *C is locally connected.* For if K denotes the set of all points in a set C satisfying (β) and (ϵ) which separate a and b in C, then it is known [See G. T. Whyburn, *Bulletin of the American Mathematical Society*, Vol. 33 (1927), p. 685; and *Transactions of the American Mathematical Society*, Vol. 32 (1930), p. 927] that $K + a + b$ is closed and compact. But we must have $K + a + b = C$; for if not, then some point x of $C - (a + b)$ fails to separate a and b, and hence by (β) cannot separate C at all; but then using (ϵ) and the well known simple chain lemma (see R. L. Moore, *loc. cit.*), we can obtain a connected chain E in C from a to b neither containing x nor having x as a limit point; and this contradicts (β), since $\bar{E}$ is then a proper subset of C.

Thus we see that the condition (α) that C be compact in the characterization [α, β, ϵ] of simple arcs given by Hausdorff (*loc. cit.*) is entirely superfluous.

† See R. L. Moore, *Bulletin of the American Mathematical Society*, abstract, Vol. 33 (1927), p. 141; K. Menger, *Monatshefte Für Mathematik und Physik*, Vol. 36 (1930), p. 210; Kuratowski, *Fundamenta Mathematicae*, Vol. 15, p. 301; and Aronszajn, *ibid.*, p. 228.

‡ See G. T. Whyburn, *American Journal of Mathematics*, Vol. 53 (1931), p. 753.

the fact that, in case N is a complete space, $\bar{H}$ is a compact locally connected continuum. Thus every two points of the space S in the generalized arcwise connectivity theorem lie together in a compact locally connected continuum M in S, and from the arcwise connectivity of M we deduce the arcwise connectivity of S.

This method of proving the arcwise connectivity theorem has additional advantages accruing from the great generality of proposition (A). In particular, (A) yields at once the celebrated theorem of Hahn-Mazurkiewicz to the effect that any compact locally connected continuum is the continuous image of an interval. Thus in a treatment of locally connected spaces the logical sequence of the results discussed above seems to be as follows: First prove theorem (A); from this we obtain as an immediate consequence the continuous image theorem of Hahn-Mazurkiewicz; next establish the arcwise connectivity theorem for compact spaces M; then with aid of (A) we obtain the generalized arcwise connectivity theorem.

6. *General Junction Theorem.* In conclusion we shall consider a very general proposition embodying a junction property of locally connected metric spaces in general from which all the above arcwise connectivity theorems follow as immediate consequences. If X is any metric space, for convenience we shall denote by X_c the space obtained by *completing* X by adding * on new points corresponding to every fundamental sequence in X.

JUNCTION THEOREM. *Every two points a and b of a connected and locally connected metric space N can be joined in N by a set H which is the image under a $(1-1)$ and doubly † uniformly continuous transformation of the set D of dyadic rational numbers ‡ on $(0,1)$. Thus H_c is a simple arc from a to b.*

This theorem is readily established with the aid of regular chains. With the aid of a theorem of the author's § it can be shown that any two points of such a space N can, for each $\epsilon > 0$, be joined in N by a simple regular ϵ-chain of locally connected sets (links) such that two successive links have just one point in common and any two links that are not successive are at a positive minimum distance apart. This established, we can proceed to construct the set $H = \sum_1^\infty P_n$ by exactly the same method as was outlined above in § 3.

---

* See Hausdorff, *Grundzüge der Mengenlehre*, (1914), p. 315.

† That is, both the transformation and its inverse are uniformly continuous.

‡ i. e., numbers of the form $m/2^n$.

§ See *American Journal of Mathematics*, Vol. 53 (1931), p. 439.

Clearly the set H_c obtained from H by *completing* it is homeomorphic with the interval $(0, 1)$ and hence is an arc in N_c from a to b. Thus we have shown that *any two points a and b of any connected and locally connected metric space N can be joined in N_c by a simple arc $H_c = ab$ such that the points H of N belonging to ab are everywhere dense on ab.* It is obvious that in case N is complete so that we have $N_c = N$, or is homeomorphic with a complete space,* we obtain the generalized arcwise connectivity theorem (see § 5). In addition to this, however, we see by virtue of the Junction Theorem that even in the general locally connected spaces, without any assumptions as to completeness or compactness, we have all the machinery necessary for the construction of simple arcs. Indeed, we can construct sets H in these spaces which, when completed, become simple arcs, and these arcs will lie in the original spaces provided these spaces are assumed to be complete.

THE JOHNS HOPKINS UNIVERSITY.

* This is known to include all G_δ subsets of any complete space. See Alexandroff, *Comptes Rendus*, Vol. 178 (1924), p. 185; also Hausdorff, *Fundamenta Mathematicae*, Vol. 5, p. 146.

BICONNECTED AND RELATED SETS.

By P. M. SWINGLE.

In Section 1 of this paper a number of simple theorems concerning biconnected * and related sets are obtained. In Section 2 various types of continua are studied with respect to the property of being the sum of mutually exclusive punctiform connected subsets. For example it is seen that a bounded indecomposable continuum can be the sum of any integral number of such sets greater than one. In the last section euclidean spaces are studied as the sum of biconnected subsets. It is shown that a necessary and sufficient condition that a euclidean space be n dimensional is that it be the sum of $n+1$, but not of n, biconnected subsets.

1. *Simple theorems concerning biconnected sets.* In this section a number of problems concerning biconnected sets are considered and a number of simple theorems proved concerning them and related sets.

It readily is seen that in order that a connected set C contain a finite dispersion set H, it is necessary and sufficient that every connected subset of C contain a finite dispersion set. Consider now the case where each of these dispersion sets contains the same number of points.

LEMMA 1. *There does not exist a connected set C, every proper connected subset of which contains the finite set H of n points, unless $n = 1$.*

For if there exists such a set C, then every proper connected subset of C contains any one point h of H. Hence C is biconnected and has the dispersion point h. As a biconnected set contains at most one dispersion point,† $n = 1$.

LEMMA 2. *If every proper connected subset of a connected set C contains a primitive dispersion set of n points, where n is any given positive*

* A biconnected set is a connected set which is not the sum of two mutually exclusive connected subsets, where it is understood that a connected set contains more than one point. These sets were defined and studied first by B. Knaster and C. Kuratowski in their interesting paper, "Sur les ensembles connexes," *Fundamenta Mathematicae*, Vol. 2, pp. 206-253. The problem of whether a biconnected set must contain a dispersion point was proposed by C. Kuratowski, *Fundamenta Mathematicae*, Vol. 3, p. 322 (19). The results of the present paper were obtained partly in consideration of this problem.

† J. R. Kline, "A Theorem Concerning Connected Sets," *Fundamenta Mathematicae*, Vol. 3, pp. 238-239.

integer, then each of these proper connected subsets contains the same primitive dispersion set H, which is also a primitive dispersion set of C. Thus n must equal one.

Let H be the primitive dispersion set of C if such exists; otherwise let $H = 0$. If $H \neq 0$, it is seen that it must contain the n points of the primitive dispersion set Q of any connected subset K of C.* And similarly Q must contain $K \times H$. Also it is seen that, if C contains a finite dispersion set D, then D contains $H \neq 0$.†

(1) Consider the case where there exists a point x of $C - H$ such that $C - x$ is connected, and so has a primitive dispersion set Q of n points. Then the dispersion set $Q + x$ contains $H \neq 0$ and so $H = Q$.

(2) Consider now the case where for each point x of $C - H$, $C - x = C_1 + C_2$ separate.‡ As $C_1 + x$ is connected H consists of $2n$ points.

Suppose that for a point h of H, $C - h = C_1 + C_2$ separate. Then H consists of $2n - 1 = 2n$ points, which is a contradiction. Hence $C - h$ is connected, and so H consists of $n + 1$ points.

Therefore $n + 1 = 2n$, and so $n = 1$. Thus H consists of the two points h_1 and h_2 say. Then for $C - x = C_1 + C_2$ separate each C_i must contain one and only one point of H. Say C_i contains h_i $(i = 1, 2)$. Let $w \neq h_1$ be a point of C_1. Then $(C_1 + x) - w = C_{11} + C_{12}$ separate, as $C - w = W_1 + W_2$ separate for each $W_i + w$ must contain points of C_1, the one which does not contain x being contained in C_1 and the other, W_2 say, containing points of C_1 as $W_2 \times C_2$ does not have w as a limit point. But as $C_{1i} + w$ is connected and only one of the primitive dispersion set can be contained in h_1, and so in H, a contradiction is obtained. Hence this case does not exist.

Therefore every connected subset of C must contain the primitive dispersion set H of C, and so by lemma 1, H consists of one point.

The truth of the following theorem and its corollary is now evident.

THEOREM 1. *In order that a connected set C contain a dispersion point it is necessary and sufficient that every proper connected subset of C contain a dispersion point.*

---

* R. L. Wilder, "On the Dispersion Sets of Connected Point Sets," *Fundamenta Mathematicae*, Vol. 6, Theorem 4, p. 216.

† R. L. Wilder, *loc. cit.*, Theorem 1, p. 214.

‡ By this notation, $C - x = C_1 + C_2$ *separate*, is meant that $C - x$ is the sum of the two non-vacuous, mutually exclusive, subsets C_1 and C_2, neither one of which contains a limit point of the other.

COROLLARY 1. *If every proper connected subset of a connected set C contain a dispersion point, then C is biconnected.*

COROLLARY 2. *There does not exist a connected set C such that, for every point x of C, the set $C - x$ is connected and has a dispersion point.*

For let K be any proper connected subset of C and let y be a point of $C - K$. Then $C - y$ contains K. Let q be the dispersion point of $C - y$. Then $(C - y) - q$ is totally disconnected and so does not contain K, which then must have q as its dispersion point. Hence as every proper connected subset of C has a dispersion point, by the above theorem, C does also, which is a contradiction.

Two theorems will be proved now concerning the dispersion point of a connected subset of a biconnected set.

THEOREM 2. *If B is biconnected and q is a dispersion point of a connected subset, K, of B, then B is the sum of two biconnected subsets, having only q in common, each of which has q as a dispersion point.**

It is evident that this theorem is true, if $B - q$ is disconnected. Consider then the case where $B - q$ is connected. Let $K - q = K_1 + K_2$ separate. Let C be the maximal connected subset of $B - K_1$, which contains the connected set $K_2 + q$. Then $B - C$ is totally disconnected as B is biconnected. Then, if $(B - C) + q = Y_1 + Y_2$ separate, where Y_1 contains the connected set $K_1 + q$, $B - C = (Y_1 - q) + Y_2$ separate. Hence $C + Y_2$ is a connected subset of $B - K_1$ and so C is not a maximal connected subset of this set. Thus $B - C + q$ must be connected. Therefore B is the sum of the two biconnected subsets C and $B - C + q$, having only q in common, which is a dispersion point of each.

THEOREM 3. *If B is biconnected and q is a dispersion point of a connected subset K of B, where K is connected im kleinen at q, then no connected subset of B has a dispersion point other than q.†*

Assume that C is a connected subset of B, having dispersion point $p \neq q$. By theorem 2, B is the sum of two biconnected subsets, mutually exclusive except for p, where each of these has p as a dispersion point. Let E be the one of these biconnected subsets which contains q. Let $E - p = E_1 + E_2$

* Whether q must be in this case a dispersion point of B itself is an unsolved problem of interest in connection with the unsolved problem, suggested by Kuratowski, of whether a biconnected set must always contain a dispersion point.

† Whether the theorem is true if K is not connected im kleinen is an unsolved problem.

separate where E_1 contains q. As K is connected im kleinen at q, K contains a connected subset N, containing q, such that $N \times (E_2 + p) = 0$. But as $E_2 + p$ is also a connected subset of B, this is impossible. Therefore $p = q$.

It readily is seen that if B is a biconnected set which is disconnected by a finite subset H, then H contains a dispersion point of B.* It follows that a biconnected set cannot contain two distinct cut points. Thus it is true that the closure † B, of a biconnected set B, contains at most one cut point, which must be a point of B. Furthermore, if this closure is a continuous curve in a euclidean space, then every point of $\bar{B}$ is contained in a simple closed curve of $\bar{B}$.‡ However it is not true necessarily that $\bar{B}$ is cyclicly connected.

It is known that there exist continuous curves, every subcontinuum of which is a continuous curve, which contain punctiform connected subsets.§ It follows from the following theorem that such continuous curves are not the sum of a countable set of arcs.

THEOREM 4. *If p is a point of a punctiform connected set P, in a euclidean space, and R is a region (sphere) containing p, then $R \times \bar{P}$ is not contained in a countable set of simple continuous arcs.*

Suppose that $R \times \bar{P}$ is contained in the countable set of arcs $A_1, A_2, \cdots$, where A_1 contains p. If x is a point of $A_i \times P \times R$, then x must be a limit point of points of P not contained in A_i, for P otherwise would have to contain a subcontinuum of A_i. Hence p is a limit point of points of $(A_2 + A_3 + \cdots) \times P \times R$, which are not contained in A_1. Consider for example the case where x_1 is a point of this set which is contained in A_5, but in no A_j where j is less than 5. Then there exists a region R_1 containing x_1, where R contains R_1, such that $R_1 \times (A_1 + A_2 + A_3 + A_4) = 0$. In a similar manner there exists a point x_2 having the same relation to x_1 that x_1 has to p, where x_2 is contained in a region R_2, such that R_1 contains $\bar{R}_2$, which contains no point common with any arc with subscript less than j where j is now greater than 5. Similarly points $x_3, x_4, \cdots$ are obtained with corresponding regions $R_3, R_4, \cdots$ where each $\bar{R}_{i+1}$ is contained in R_i and has points common with fewer arcs than R_i did. Hence it is seen that the set $x_1, x_2, \cdots$ has a limit point in $\bar{P}$ which none of the arcs $A_1, A_2, \cdots$ contains.

---

* R. L. Wilder, *loc. cit.*, Theorems 1, 4, and 10.

† By the closure of a set B is meant the set $\bar{B}$ consisting of B and the limit point of B.

‡ This follows from Theorem 11 of W. L. Ayres' paper "Concerning Continuous Curves in Metric Space," *American Journal of Mathematics*, Vol. 51 (1929), p. 591.

§ B. Knaster and C. Kuratowski, "A Connected and Connected im Kleinen Point Set Which Contains No Perfect Subset," *Bulletin of the American Mathematical Society*, Vol. 33 (1927), pp. 106-109.

2. *Continua as the sum of punctiform connected subsets.* The following lemma can be proved readily by means of transfinite induction and Zermelo's postulate,* as it is possible to continue to choose points for the K_f's from the C_i's and for the Q_g's from the P_i's, since the product of the power of the set of points chosen and $m+2$ is less than n, except where this power is itself n.

LEMMA A. *Let h be the first ordinal number of power m and b be the first ordinal number of power n, where $mn=n$ and $m<n$. Let (C_i) and (P_i) $(i=1,2,\cdots b)$ be two classes of point sets, each C_i and each P_i being of power n. Then (1) there exists a class (K_f) $(f=1,2,\cdots,h)$, of power n, of mutually exclusive point sets K_f such that each K_f contains a point of every C_i of (C_i) and (2) there exists a class (Q_g) $(g=1,2)$ of mutually exclusive point sets Q_g such that each Q_g contains a point of every P_i of (P_i); also for each K_f and Q_g, $K_f \times Q_g = 0$.*

THEOREM A. *Let W be a bounded indecomposable continuum, lying in a euclidean space, and let m be any cardinal number greater than one and less than c, the power of the linear continuum. Then W is the sum of m mutually exclusive, punctiform, connected subsets, each of which is dense in W.*†

For let (C) be the class of the continua of our space, each of which separate points of W, and let (P) be the class of the perfect sets of W. Both (C) and (P) have power c. Hence by Zermelo's well-ordering theorem ‡ these sets can each be well-ordered obtaining classes (C_i) and (P_i) $(i=1,2,\cdots,k)$, where the power of k is c.

Let a and b be any two points of W separated by C_i of (C_i) and let T be any *composant* § of W. Then as $T+a+b$ is connected, C_i must contain

---

* W. Sierpinski, *Leçons sur les Nombres Transfinis* (1928), pp. 164-165, p. 212, and p. 137.

† The method of proof used here is due to B. Knaster and C. Kuratowski, *Fundamenta Mathematicae*, Vol. 2, pp. 250-251. It is to be noted that if any of the above punctiform, connected subsets of W is a biconnected set B, then the problem concerning the existence of a dispersion point is solved in the negative. For if B is an indecomposable continuum, B cannot contain a cut-point. See P. M. Swingle, "Two Types of Connected Point Sets," *Bulletin of the American Mathematical Society*, Vol. 37 (1931).

‡ E. Zermelo, "Beweis dass jede Menge wohlgeordnet werden kann," *Mathematische Annalen*, Vol. 59 (1904), pp. 514-516.

§ For definitions and properties see Z. Janiszewski and C. Kuratowski, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 1, pp. 217-222.

a point of $T + a + b$. As there exist c composants * of W, it follows that $C_i \times W$ is of power c. Also it is known that P_i is of power c.

Since $c \times c = c$, it follows that $m \times c = c$. Then by lemma A there exist mutually exclusive sets $K_1, K_2, \cdots, K_h$, where the power of h is m, and mutually exclusive sets Q_1 and Q_2 of W, where (1) $K_f \times Q_g = 0$ $(f = 1, 2, \cdots, h;\ g = 1, 2)$, (2) each K_f contains a point of every C_i of (C_i), and each Q_g contains a point of every P_i of (P_i).

Let all the points of W, which are not contained either in a K_f or in Q_2, be added to K_1, obtaining a new set K. Then each of the sets K, $K_2 + Q_2$, K_3, $K_4, \cdots, K_h$ is punctiform, for Q_1 contains a point, which is contained in none of these sets except K, of each perfect set of W; and Q_2 contains a point, which is not contained in K, of each perfect set of W. As each of these sets contains a point of every C_i of (C_i), they are connected.† And it is evident that each of these sets is dense in W. Therefore it is seen that the theorem is true.

Examples of indecomposable continua, every proper subcontinuum of which is a simple continuous arc, are well known. Hence there exist indecomposable continua no proper subcontinuum of which contains a punctiform connected subset. For such continua the following corollary is seen to be true.

COROLLARY A. *Let W be a bounded indecomposable continuum, lying in a euclidean space, no proper subcontinuum of which contains a punctiform connected subset. Let m be any cardinal number greater than one and less than c, the power of the linear continuum. Then W is the sum of m mutually exclusive, widely connected subsets, each of which is dense in W.*‡

By a proof similar to the above it is seen that *a bounded domain D, lying in a euclidean space, is the sum of m mutually exclusive, punctiform, connected subsets, each dense in D, where m is any cardinal number greater than one and less than c, the power of the linear continuum.*§

A domain D, lying in a euclidean space of n dimensions, may be the sum of m biconnected subsets, which are mutually exclusive except for one point and are each dense in D, where m is any cardinal number greater than n and less than c, the power of the linear continuum. For consider the following example.

---

* S. Mazurkiewicz, "Sur les continus indécomposables," *Fundamenta Mathematicae*, Vol. 10, pp. 305-310.

† B. Knaster and C. Kuratowski, *Fundamenta Mathematicae*, Vol. 2, pp. 233-234, Theorem 37.

‡ A connected set will be said to be *widely connected* if every proper connected subset is dense in it. See P. M. Swingle, "Two Types of Connected Sets," *loc. cit.*

§ See F. Bernstein, *Leipziger Berichte*, Vol. 60 (1908).

Example A. Consider in the euclidean plane the straight line interval bc and the point a not on the line bc. It will be shown that the domain bounded by the triangle abc is the sum of three biconnected sets, having only a in common, which is a dispersion point of each. The interval bc is the sum of three mutually exclusive subsets U_1, U_2, and U_3, each of power c and each dense in bc. Let W_i $(i = 1, 2, 3)$ be the set of straight line intervals each of which join a to a point of U_i. Then by a proof, similar to that of theorem A, it is seen that W_i is the sum of two punctiform connected subsets, B_{1i} and B_{2i}, having only a in common and each dense in W_i. But as a is a dispersion point of B_{ji} $(j = 1, 2)$, B_{ji} is biconnected. Hence the interior of the triangle plus its boundary is the sum of the three biconnected subsets, $B_{11} + B_{12}$ $B_{21} + B_{23}$, and $B_{22} + B_{13}$, having only a in common, a being the dispersion point of each of these sets, and each of these sets is dense in the interior of the triangle. It is seen readily that the above method can be modified to show that the interior of the triangle abc is the sum of three biconnected subsets, having only one point in common, where each of these biconnected subsets is contained in the interior of the triangle. And the above method easily is modified to show that either of these sets is the sum of m such biconnected subsets, even if $n > 2$.

THEOREM 5. *An n-dimensional compact metric space, S, is never contained in the sum of n biconnected sets.*

Assume that S is contained in the sum of the n biconnected sets B_i $(i = 1, 2, \cdots, n)$. As a biconnected set does not contain two distinct connected subsets, each set of $n + 1$ mutually exclusive n-dimensional spheres of S contains one such sphere, S_n say, which does not contain a connected subset of any B_i. But S_n contains an n-dimensional Cantorian multiplicity C_n.* Say for example that B_n contains two distinct points of C_n, as one of the B_i's will. Then there exists a subcontinuum of S_n, and so a closed subset K_{n-1} of $C_n - B_n$, which separates B_n.† And as C_n is a Cantorian multiplicity,‡ K_{n-1} must be of dimension $n - 1$. But as a closed subset of a compact space is itself a compact space, K_{n-1} contains an $(n - 1)$-dimensional Cantorian multiplicity C_{n-1} as a subset, having the property that it is contained in $n - 1$ of the biconnected sets B_i. Hence, proceeding as above, by induction it follows that there exists a one-dimensional Cantorian multi-

* K. Menger, *Dimensiontheorie* (1928), p. 217, Theorem 8.

† C. Kuratowski and B. Knaster, "Sur les ensembles connexes," *Fundamenta Mathematicae*, Vol. 2, pp. 233-235, Theorem 37.

‡ P. Urysohn, "Mémoire sur les multiplicités Cantoriennes," *Fundamenta Mathematicae*, Vol. 7, p. 124, Definition.

plicity C_1, which is contained in one of the biconnected sets B_i. Hence, as C_1 is a continuum contained in a punctiform set, a contradiction has been obtained.

COROLLARY 3. *If D is a domain of a compact metric space containing a continuum C, which is the sum of two biconnected sets, then $\overline{D - C} = \bar{D}$.*

This is true because no point of C is an interior point of C.

It has been shown that a domain in the plane is not the sum of two biconnected sets. It will be shown now that there is a continuum in the plane, which is the sum of two mutually exclusive biconnected sets. Let, in the euclidean plane, T be a Cantor ternary set on the straight line interval bc, and let a and d be two points not on the line bc, where a and d lie on opposite sides of this line. Let W be the continuum obtained by joining the points of T to each of the points a and d by straight line intervals. Then $W - d$ is seen to be, by a proof similar to that of theorem A, the sum of two biconnected sets B_1 and B_2, having a in common. Therefore W is the sum of the two mutually exclusive biconnected subsets B_1 and $B_2 - a + d$. It is to be noted that W is a compact metric space of dimension one. Also each of the biconnected subsets is of dimension one.

THEOREM 6. *An n-dimensional euclidean space, E_n, does not contain an n-dimensional biconnected subset.*

Assume that E_n contains an n-dimensional biconnected subset B. Then by definition of an n-dimensional space, there exists an n-dimensional sphere S_n whose boundary S_{n-1} contains an $n-1$ dimensional punctiform subset of B. And similarly S_{n-k} $(k = 1, 2, \cdots, n-2)$ will have as boundary S_{n-k-1}, containing an $n-k-1$ dimensional punctiform subset of B. Hence by induction there exists a one-dimensional sphere S_1, which contains a punctiform one-dimensional subset of B. Thus a contradiction is obtained.*

It readily is seen that, if n is greater than one, an n-dimensional euclidean space may contain a k-dimensional biconnected subset, where k is any integer from one to $n-1$.

3. *Euclidean spaces as the sum of biconnected sets.* In this section it will be shown that a euclidean space can be the sum of mutually exclusive biconnected subsets.†

---

* P. Urysohn, *loc. cit.*, p. 76.

† See P. M. Swingle, "Generalizations of Biconnected Sets," *American Journal of Mathematics*, Vol. 53 (1931), Problem 2, p. 392. Parts of this problem are not answered in this section; for example, whether a euclidean plane can be the sum of c mutually exclusive biconnected subsets, where c is the power of the linear continuum, remains unanswered.

In Example A of the previous section it was shown, by means of Zermelo's postulate, that, if T was the interior of a triangle *bae* in E_2, then $\bar{T}$ is the sum of m biconnected subsets, where m is any cardinal number greater than 2 but less than c, mutually exclusive except for one point which is a dispersion point. Each of these biconnected subsets are dense in T.

It will be shown now that *the euclidean plane E_2 is the sum of m mutually exclusive biconnected subsets, where m is any integer greater than* 2. For let S be the points of E_2, and let e be the center of a regular polygon P of m sides. Let on the circle, with center e and radius one, U_j $(j = 1, 2, 3)$ be three mutually exclusive punctiform point sets of power c and each dense on this circle, whose sum is this circle. Let W_j be the points on the straight line intervals, from e to infinity, passing from e through the points of U_j. Then by a proof, similar to that of theorem A, it is seen that S is the sum of m biconnected subsets, $G_1, G_2, \cdots, G_m$, mutually exclusive except for the common dispersion point e and each dense in S.

Let a_i $(i = 1, 2, \cdots, m)$ be the vertices of P. Each triangle $ea_{i+1}a_i$, where $m + 1 = 1$, is the sum of m biconnected subsets $B_{(i+1)g} + a_{i+1}$ $(g = 1, 2, \cdots, m)$, where the $B_{(i+1)g}$'s are mutually exclusive and the point a_{i+1} is the dispersion point of each biconnected subset. Furthermore $(B_{(i+1)(i+1)} + a_{i+1}) \times ea_i$ can be taken as vacuous, while each of the other biconnected sets is dense on ea_i. Let $G_{(i+1)e}$ be the points of G_{i+1} exterior to P and let

$$N_{i+1} = B_{(i+1)1} + B_{(i+1)2} + \cdots + B_{(i+1)m} + a_{i+1} + G_{(i+1)e}.$$

Then N_{i+1} is a biconnected set with dispersion point a_{i+1} and dense in S. Hence the euclidean space E_2 is the sum of the m mutually exclusive biconnected subsets N_{i+1}.

That E_n, $n > 2$, *is the sum of m mutually exclusive biconnected subsets, where m is any integer greater than n,* is shown in a similar manner.

A similar proof will show that E_n *is the sum of a countable infinity of mutually exclusive biconnected subsets.* For consider E_2. Take the regular polygon with five sides. Let $a_6, a_7, \cdots$ be an infinite sequence of distinct points on a_5a_1, with sequential limit point a_1. The proof then follows as above.

That E_3 *is the sum of c mutually biconnected subsets, where c is the power of the linear continuum,* follows from the fact that E_3 is the sum of c mutually exclusive planes, each of which is the sum of a finite number of biconnected subsets.

That E_2 is the sum of a countable infinity of mutually exclusive biconnected subsets, each with a dispersion point but not each dense in E_2, can be shown by means of the following interesting example of a biconnected set.

In E_2 take a cartesian coördinate system. By a biconnected set

$$((i,j)(i,j+1)(i+1,j)),$$

or similar set, will be meant one dense, and contained, in a triangle having these points as vertices and with the point $(i,j+1)$ as dispersion point of this biconnected set. It will be further understood that this biconnected set does not contain any point, excepting its dispersion point, on the line interval from (i,j) to $(i,j+1)$ nor on the line interval from $(i+1,j)$ to $(i,j+1)$, but that it is dense on the interval from (i,j) to $(i+1,j)$, excepting that the biconnected set $((0,0)(0,1)(1,0))$ also has nothing common with the interval from $(0,0)$ to $(1,0)$. Let i and j take on all values ± 1, ± 2, $\pm 3, \cdots, 0$.

Let $T_{(+j)}$ $(j=1,2,\cdots)$ be the set of all the points, excepting the dispersion point, of the following infinite set of biconnected sets:

$$\begin{aligned}
&((i,j)(i+1,j)(i+1,j-1)) && (i=0,1,2,\cdots),\\
&((i,j)(i-1,j)(i,j-1)) && (i=0,-1,-2,\cdots),\\
&((i,j-1)(i+1,j-1)(i,j)) && (i=1,2,\cdots),\\
&((i,j-1)(i-1,j-1)(i-1,j)) && (i=0,-1,-2,\cdots)
\end{aligned}$$

and

$$((0,j-1)(0,j)(1,j-1)).$$

To obtain $T_{(-j)}$ $(j=1,2,\cdots)$ rotate $T_{(+j)}$ 180° about the X-axis and translate the resulting set, parallel to itself, through an X distance of -1; for example $((0,0)(0,1)(1,0))$ will become $((-1,0)(-1,-1)(0,0))$. However let this last set be dense on the interval from $(-1,0)$ to $(0,0)$.

Let g be the point $(0,1)$. Then

$$g+T_1+T_2+\cdots+T_{-1}+T_{-2}+\cdots=B$$

is a biconnected set everywhere dense in E_2, connected im kleinen at g and having g as dispersion point. It has the interesting property that only one of the quasi-components of $B-g$ contains a point outside of the triangle $((0,0)(0,1)(1,0))$.

It is easy to see that if S is the set of all the points of E_2, excepting the countable infinity of points of E_2 which are dispersion points of the above biconnected sets, and if $g_1=(0,2)$ and $g_2=(0,3)$, then $S+g+g_1+g_2$ is the sum of three mutually exclusive biconnected sets, each dense in E_2 and having dispersion points g, g_1, and g_2 respectively. And it is further seen that E_2 is the sum of a countable infinity of mutually exclusive biconnected sets, three of which are dense in E_2 and the remaining ones are each con-

tained in one of the triangles used above. Similar results can be obtained for E_n, $n > 2$.

From the results of the last two sections the following theorem now will be proved.

THEOREM B. *In order that a euclidean space E be n dimensional, n a positive integer greater than one, it is necessary and sufficient that E be the sum of $n+1$ but not of n biconnected subsets.**

The condition is necessary. For by theorem 5, E_n is not the sum of n biconnected subsets. And in this section it is shown, by means of Zermelo's postulate, that E_n is the sum of $n+1$ mutually exclusive biconnected subsets.

The condition is sufficient. For E is the sum of $n+1$ biconnected subsets by hypothesis, and so is not k dimensional, k greater than n, for E_k contains an E_{n+1}, which by theorem 5 is not contained in the sum of $n+1$ biconnected sets. And it is not a k dimensional euclidean space, k less than n, for such a space could be the sum of n biconnected subsets, excepting for $k=1$ which would not contain a biconnected set. Hence E must be n dimensional.

OHIO STATE UNIVERSITY.

* For an interesting similar result see P. Urysohn, *Fundamenta Mathematicae*, Vol. 8, Corollary, p. 345; see also Karl Menger's *Dimensiontheorie*, p. 113 and p. 120

A NOTE ON SPACES HAVING THE S PROPERTY.

By GORDON T. WHYBURN.

1. A metric space D is said to have property S * provided that for each $\epsilon > 0$, D is the sum of a finite number of connected sets each of diameter $< \epsilon$. Since all such spaces have only a finite number of components, little generality is lost by supposing them connected. If N is any subset of a metric space D and ϵ is any positive number, we shall denote by $T_\epsilon(N)$ the set of all points x of D which can be joined to some point p of N by a chain of connected subsets $L_1, L_2, \cdots, L_n$ of D such that for each i, $\delta(L_i) < \epsilon/2^i$, $L_1 \supset p$, $L_n \supset x$, and any two successive sets (links) L_i and L_{i+1} have at least one common point. Such a chain will be called a chain of type T_ϵ or simply type T.

2. THEOREM. *If the metric space D has property S, then every set $T_\epsilon(N)$ has property S.*

For let δ be any positive number, and let us choose an integer k such that $\sum_k^\infty \epsilon/2^i < \delta/4$. Let E be the set of all points in $T_\epsilon(N)$ which can be joined to N (i. e., to some point of N) by a chain of type T which has at most k links. Let us express D as the sum of a finite number of connected sets each of diameter $< \epsilon/2^{k+1}$; and of these sets, let $Q_1, Q_2, \cdots, Q_n$ be the ones which contain at least one point of E. Then we have $E \subset \sum_1^n Q_i$. Now for each i, $Q_i \subset T_\epsilon(N)$; for Q_i contains a point x of E, and x can be joined to N by a chain $L_1, L_2, \cdots, L_r$ of type T having k links or less; and since $\delta(Q_i) < \epsilon/2^{k+1}$, therefore $[L_1, L_2, \cdots, L_r, Q_i]$ is a chain of type T, and hence $Q_i \subset T_\epsilon(N)$. For each i, $1 \leqq i \leqq n$, let W_i be the set of all points of $T_\epsilon(N)$ which can be joined to some point of Q_i by a connected subset of $T_\epsilon(N)$ of diameter $< \delta/4$. Then for each i, W_i is a connected subset of $T_\epsilon(N)$ of diameter $< \delta$. It remains only to show that $T_\epsilon(N) \subset \sum_1^n W_i$. To this end let x be any point of $T_\epsilon(N)$ and let $L_1, L_2, \cdots, L_m$ be a chain of type T joining x to N. Obviously we need only consider the case in which x does not belong to E, and in this case $m > k$. Then since $L_k \subset E$, it

* See Sierpinski, *Fundamenta Mathematicae*, Vol. 1, p. 44; and R. L. Moore, *ibid.*, Vol. 3, p. 232.

follows that for some j, $L_k \cdot Q_j \neq 0$; and since $\sum_k^\infty \epsilon/2^i < \delta/4$, it follows that $\delta(\sum_k^m L_i) \leq \sum_k^m \delta(L_i) < \delta/4$. Hence $\sum_k^m L_i$ is a connected subset of $T_\epsilon(N)$ of diameter $< \delta/4$ which joins x to a point of Q_j. Therefore $x \subset W_j$, and our theorem is proved.

3. COROLLARY 1. *Any metric space D having property S is the sum of a finite number of arbitrarily small connected subsets each having property S. Furthermore these subsets may be chosen either as open sets or as closed sets.*

For let δ be any positive number, let $\epsilon = \delta/3$, and let $D = \sum_1^n D_i$, where each D_i is connected and of diameter $< \epsilon$. Then, for each i, $T_\epsilon(D_i)$ is connected and of diameter $< \delta$ and clearly $D = \sum_1^n T_\epsilon(D_i)$. Now the sets $[T_\epsilon(D_i)]$ themselves are open; and since it is true that if a set E has property S, so does every set E_0 such that $E \subset E_0 \subset \bar{E}$, it follows that the sets $[\overline{T_\epsilon(D_i)}]$ have property S, and of course they are closed.

From this corollary it follows that in any metric space having property S there exists a monotone decreasing *fine* subdivision into connected sets. In other words, we can subdivide such a space into a finite number of connected sets each having property S and being of diameter < 1; then we can subdivide each of these sets into a finite number of connected sets of diameter $< 1/2$, and so on indefinitely.

4. COROLLARY 2. *Any point p of a metric space D having property S is contained in an arbitrarily small connected open set (region) which has property S.*

To see this we have only to take $N = p$, and then the set $T_\epsilon(p)$ is the desired region. Since $\overline{T_\epsilon(p)}$ also has property S and hence is locally connected (because any set having property S is locally connected), we have shown that p is contained in an arbitrarily small region whose closure is locally connected.

5. For any point p of a metric space D and for each $\epsilon > 0$ let $R_\epsilon(p)$ be the set of all points x of D which can be joined to p by a chain of connected sets $L_1, L_2, \cdots, L_n$ in D such that $L_1 \supset p$, $L_n \supset x$, any two successive links have a common point, and for each i, $\delta(L_i) < \epsilon - \sum_1^i \delta(L_j)$.* Then by

* See my paper "Concerning S-Regions in Locally Connected Continua," appearing in *Fundamenta Mathematicae*.

essentially the same argument as given in § 2 it is shown that if D has property S, every set $R_\epsilon(p)$ likewise has property S.

Now let us suppose D has property S. For any ordered pair of points x, y in D let us denote by $\rho^+(x, y)$ the greatest lower bound for the aggregate of numbers $[\epsilon]$ such that $y \subset R_\epsilon(x)$, and let D^+ denote the space whose points are the same as those of D but in which distances are defined by the function ρ^+. It is readily shown that ρ^+ satisfies all the axioms for a metric except the symmetry axiom $[\rho(x, y) = \rho(y, x)]$, which may not be satisfied; and thus D^+ is a quasi metric space. Furthermore, D^+ is homeomorphic with D; and the spherical neighborhoods in D^+ are exactly the sets $R_\epsilon(p)$ (thus they are connected and have property S). This naturally raises the question of the existence of *some* distance function ρ which satisfies *all* the axioms for a metric and which gives these properties to the new space. In other words, *Is every metric space D which has property S homeomorphic with some metric space D' in which all spherical neighborhoods are connected and have property S?*

A positive solution to this problem even in the case of compact connected spaces D having property S (i. e., compact locally connected continua) would yield a very convenient and useful machinery with which to study such spaces.

THE JOHNS HOPKINS UNIVERSITY.

REMARKS ON THE RECURRENT BEHAVIOR OF THE CHARACTERISTICS ON A TORUS.

By AUREL WINTNER.

A paper of the author on the Levi-Civita problem of mean motion (to appear in the *Annali di Matematica*) seems to lead to some contradictions with an important discovery by Denjoy, recently published in the *Comptes Rendus*. On combining the theorem of Denjoy with other results it will be shown in the present note that, in reality, those apparent contradictions do not exist at all.

Let $f(\vartheta, t)$ be a real-valued continuous function defined in the real (ϑ, t)-plane and possessing the period 1 with respect to both ϑ and t. Let T denote the closed orientable surface of genus one representing the fundamental domain $0 \leqq \vartheta \leqq 1$, $0 \leqq t \leqq 1$. The function f should possess the additional property that the differential equation

$$d\vartheta/dt = f(\vartheta, t) \tag{1}$$

admits, for a given real initial condition $\vartheta(t_0) = \vartheta_0$, only one solution $\vartheta = \vartheta(t)$, depending on the initial condition in a continuous manner. This additional condition which will be designated as the condition Γ is, for instance, fulfilled if $\partial f/\partial\vartheta$ exists and is continuous. If every solution

$$\vartheta = \vartheta(t)\,; \quad -\infty < t < +\infty \tag{2}$$

of (1) represents an everywhere-dense point-set on the torus T we shall say that the function f is of the ergodic type.

It is known since Poincaré* that if f fulfills the condition Γ then there exists for every solution (2) of (1) a constant μ and a *bounded* function $\omega(t)$ so that

$$\vartheta(t) = \mu \cdot t + \omega(t) \tag{3}$$

where the "mean motion" or "rotation number" μ is independent of the special choice of the integration constant $\vartheta(t_0) = \vartheta_0$ and is, therefore,

* Cf. H. Kneser, "Reguläre Kurvenscharen auf den Ringflächen," *Mathematische Annalen*, Vol. 91 (1924), pp. 135-154, where references are also given. The proof of a lemma occurring in the demonstration has been simplified by J. Nielsen, "On topologiske Afbildninger af en Jordankurve paa sig selv," *Matematisk Tidsskrift*, B, No. 3 (1928), p. 39-46. A very simple direct proof of (3) has recently been given by A. Weil, "On Systems of Curves on a Ring-Shaped Surface," *The Journal of the Indian Mathematical Society*, Vol. XIX (1931), pp. 109-114.

uniquely determined by the double-periodic function f. It may be mentioned that the number μ which is obtained by an infinite process of iteration cannot be less than the minimum or larger than the maximum of the double-periodic function $f(\vartheta, t)$. In fact, on substituting (3) in (1) and using the fact that $\omega(t)$ is a bounded function, we obtain that the time-average

$$\lim_{t=+\infty} \int_0^t f(\vartheta(\tau), \tau)\, d\tau/t$$

exists and is equal to μ.

On account of the boundedness of the remainder term $\omega(t)$ one may expect that f is of the ergodic type provided that μ is irrational. This is, however, not necessarily the case, even if $\partial f/\partial\vartheta$ exists and is continuous. On refining the methods of his earlier paper on this problem Kneser has shown in a note to appear in the *Annali di Matematica* that if the condition Γ is fulfilled and if μ is irrational the function f is then and only then of the ergodic type if the remainder term $\omega(t)$ is, for all values of the integration constant $\vartheta(t_0) = \vartheta_0$, an almost-periodic function in the sense of Bohr. On the other hand, it has been conjectured by Poincaré (1885) that f is necessarily of the ergodic type if it is regular analytic in the real domain (provided that μ is irrational). The correctness of this conjecture which is of great importance for the dynamics of Birkhoff * has been proved by Denjoy † in his fundamental note mentioned above. Denjoy shows that even the condition of analyticity is superfluous, viz., it is sufficient that a certain "real" condition is fulfilled which is only a little stronger than the non-sufficient condition Γ.

We combine now these results with the following fact, explicitly pointed out also by Bohl:‡ if the condition Γ is fulfilled and μ is irrational the function f is then and only then of the ergodic type if there exists, for any value of the integration constant $\vartheta(t_0) = \vartheta_0$, a continuous double-periodic function

$$\rho = \rho(\xi, \eta) = \rho(\xi + 1, \eta) = \rho(\xi, \eta + 1),$$
$$-\infty < \xi < +\infty, \quad -\infty < \eta < +\infty$$

for which

---

* Cf., for instance, G. D. Birkhoff, "Surface Transformations and their Dynamical Applications," *Acta Mathematica*, Vol. 43 (1920), p. 79 etc., and "On the Periodic Motions of Dynamical Systems," *Acta Mathematica*, Vol. 50 (1927), pp. 359-379.

† A. Denjoy, "Sur les caractéristiques à la surface du tore," *Comptes Rendus des Séances de l'Académie des Sciences*, Vol. 194 (1932), pp. 830-833.

‡ P. Bohl, "Ueber die hinsichtlich der unabhängigen und abhängigen Variabeln periodische Differentialgleichung erster Ordnung," *Acta Mathematica*, Vol. 40 (1916), p. 321 etc.

$$\omega(t) = \rho(\mu \cdot t, t). \tag{4}$$

Since any function $\omega(t)$ possessing a representation (4) is almost-periodic in the sense of Bohr * it follows from the result of Kneser mentioned above that if the remainder term $\omega(t)$ is at all almost-periodic then it possesses only two linearly independent frequencies (viz., μ and 1). On the other hand, on combining the remark of Bohl with the theorem of Denjoy it follows that $\omega(t)$ is in the analytic case necessarily almost-periodic, provided μ is irrational.

Let $a(t)$, $b(t)$, $c(t)$, $d(t)$ be four functions which are real-valued and regular analytic for $-\infty < t < +\infty$ and possess the period 1. On denoting by

$$x = x(t),\ y = y(t);\quad x(t)^2 + y(t)^2 > 0 \tag{5}$$

a real solution of the differential equations

$$dx/dt = a(t)x + b(t)y,\quad dy/dt = c(t)x + d(t)y \tag{6}$$

and placing

$$x(t) = r(t)\cos(\vartheta/2\pi),\quad y(t) = r(t)\sin(\vartheta/2\pi) \text{ where } r(t) > 0 \tag{7}$$

one obtains from (6), according to Levi-Civita,† the single differential equation

$$\begin{aligned} d\vartheta/dt = 2\pi\{&c(t)\cos^2(\vartheta/2\pi) \\ &+ [d(t) - a(t)]\cos(\vartheta/2\pi)\sin(\vartheta/2\pi) - b(t)\sin^2(\vartheta/2\pi)\} \end{aligned} \tag{8}$$

which is, of course, a special case of (1) and satisfies the regularity condition of Denjoy. In the note mentioned above I pointed out that $\omega(t)$ is, in the special case (8) of (1), not necessarily almost-periodic. It follows, therefore, from the theorem emphasized at the end of the previous paragraph that μ cannot be, in these cases, an irrational number. On the other hand, in the presentations of the problem ‡ it is usually shown that $\omega(t)$ is, even under the single condition Γ, not only almost-periodic but also periodic if μ is a rational number. Accordingly $\omega(t)$ is then in the case (8) *always* almost-periodic. This is, however, in contradiction with the following examples:

Suppose that the four coefficients of (6) are not only periodic but even

---

* H. Bohr, "Zur Theorie der fastperiodischen Funktionen. II.," *Acta Mathematica*, Vol. 46 (1925), p. 134 etc.

† T. Levi-Civita, "Sur les équations à coefficients périodiques et sur le mouvement moyen du noeud lunaire," *Annales de l'École Normale Supérieur* (3), Vol. 28, pp. 325-376.

‡ Cf., for instance, A. Weil, *loc. cit.*

independent of t. Equation (8) is also in this case a special case of equation (1). Suppose further that the characteristic numbers of the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

are real and that the determinant of the matrix is positive. There are three cases possible, according as the characteristic numbers and the elementary divisors are simple or double. However, in all three possible cases the origin of the (x, y)-plane is a node (noeud, Knoten),* i. e., the curve (5) which is in general not a straight line possesses in the point $x = y = 0$ a tangent. We conclude from (7) that the function $\vartheta(t)$ possesses for † $\lim t = \infty$ a finite limit α but is not equal to this constant α for all values of t (provided the curve is not a straight line which may be excluded). Since the limit α exists it follows from (3) that $\vartheta(t)$ coincides with the bounded function $\omega(t)$. Furthermore, since $\vartheta(t)$ is not constant and possesses for $\lim t = \infty$ a limit it follows that $\omega(t)$ is *not* an almost-periodic function.

This paradox is cleared up as follows: On reading over the usual demonstrations for the fact that $\omega(t)$ is a periodic function for rational values of μ one may observe that the proof breaks down if the rational rotation number μ is equal to zero as is the case in our examples.

THE JOHNS HOPKINS UNIVERSITY.

* Cf., for instance, E. Kamke, "Differentialgleichungen reeller Funktionen," Leipzig, 1930, p. 200 etc. Intuitive pictures for the different kinds of nodes are given by O. Perron, "Ueber die Gestalt der Integralkurven einer Differentialgleichung erster Ordnung in der Umgebung eines singulären Punktes. (Erster Teil)," *Mathematische Zeitschrift*, Vol. 15 (1922), p. 123.

† It depends on the matrix of the coefficients whether $\infty = +\infty$ or $\infty = -\infty$.

SOME REMARKS ON THE JOIN OF TWO COMPLEXES AND ON INVARIANT SUBSETS OF A COMPLEX.

By EGBERT R. VAN KAMPEN.

In this paper we consider some miscellaneous properties of combinatorial topology mostly found during the preparation of a course on this subject. The first three sections deal with the homology characters of a join; it is found useful to introduce homology characters in a combinatorial way for a kind of open complex slightly more general than is customary. After that we prove some simple properties of invariant subsets of complexes, soon concentrating on the problem of dividing any complex into the smallest invariant subsets that can be found with the methods now at the disposal of combinatorial topology. To do this we have to prove a number of properties on the invariance in a complex of the invariants of the neighborhood complexes of the different simplexes of the complex, using some of the properties of joins proved previously.

1. As the cells of the join $(K_1 \cdot K_2)$ and the product $(K_1 \times K_2)$ of two complexes (K_1) and (K_2) as well as their incidence relations are completely determined by K_1 and K_2 it is a matter of calculation to find their homology characters when K_1 and K_2 are given. Superficially the computation seems easier for the product because in the product there is one $(p+q)$-cell corresponding to every pair formed by a p-cell of K_1 and a q-cell of K_2 and no other cell; the boundary of the $(p+q)$-cell corresponds to all the pairs which can be formed by replacing the p- or the q-cell by its boundary cells. The structure of a join is slightly more complicated. This induced A. B. Brown * to reduce the computation for $K_1 \cdot K_2$ to that for $K_1 \times K_2$ by means of a relatively complicated reasoning on different kinds of chains on the join; he does not find any result on torsion numbers. However, when we introduce formally a (-1)-dimensional simplex into all the relevant complexes forming the boundary of all vertices and such that its join with any cell is that cell, then there is in the join a $(p+q+1)$-simplex corresponding to every pair formed by a p-simplex of K_1 and a q-simplex of K_2; with analogous boundary relations as for the product. The result for the product † can now be transcribed for the join without any additional calculation.

* "On the Join of two Complexes," *Bulletin of the American Mathematical Society*, Vol. 37 (1931), p. 417.

† See for instance E. R. van Kampen, "Die kombinatorische Topologie und die Dualitätssätze," Van Stockum, *Den Haag*, p. 37.

THEOREM I. *Suppose the homology characters of K_1, K_2 and $K_1 \cdot K_2 = K$ to be computed after the introduction of a (-1)-simplex which is on the boundary of every 1-simplex (that means putting R° equal to the number of components minus one), then we have for the homology characters of K:*

$$R^p(K) = \sum_{k+l+1=p} R^k(K_1) \cdot R^l(K_2).$$

The numbers $\tau_i^p(K)$ are the elementary divisors of a matrix of which the elements outside of the diagonal are zero while the elements in the diagonal are:

$R^k(K_2)$ *times the numbers* $\tau_i^{p-k-1}(K_1)$, $(k = 0, \cdots, p-1)$;
$R^k(K_1)$ *times the numbers* $\tau_i^{p-k-1}(K_2)$, $(k = 0, \cdots, p-1)$;

the greatest common divisors of all pairs

$\tau_i^k(K_1)$, $\tau_j^{p-k-1}(K_2)$, $(k = 1, \cdots, p-2)$,
and of all pairs $\tau_i^k(K_1)$, $\tau_j^{p-k-2}(K_2)$, $(k = 1, \cdots, p-3)$.

2. We are now going to define homology characters of L modulo M, where L and M are sums of the interiors of certain simplexes of a complex K, L containing M, and M containing its limit-points insofar as these belong to L.* As the case considered here is barely more general than that worked out in detail by S. Lefschetz in his Topology, we give only the definitions, none of the proofs.

The homology characters of L modulo M are defined as the homology characters of $\bar{L}$ modulo $\bar{M}$ where $\bar{L}$ is a subcomplex of a subdivision of K approximating L and $\bar{M}$ is the intersection of $\bar{L}$ and M, which is also a complex. For $\bar{L}$ we take the point set obtained from K by applying the following process to every simplex of K of which the interior is not on L: Construct a copy U of the neighborhood complex of that simplex in a very small neighborhood of one of its interior points and take away from K the join of U and the simplex, except the join of U and the boundary of the simplex. U must be taken so small, that the parts taken away for two simplexes do not have a common point when the simplexes are not incident. $\bar{L}$ is a subcomplex of a subdivision of K; any chain on L of which some part is on M can be deformed into a chain of $\bar{L}$, the part on M remaining on M and finally moving into $\bar{M}$. The homology characters of L modulo M are invariant under transformation of K into another complex carrying L and M into sums of interiors of simplexes. This follows from the following:

* If this is not true from the beginning the logical approximation for M is exactly the inclusion of those limit-points.

THEOREM 2. *The homology characters of L modulo M can be defined by means of singular chains on L neglecting those contained in M.*

3. We want to find point-sets L_3 and M_3 on $K_1 \cdot K_2 = K_3$ such that the formulas of theorem 1 give the homology characters of L_3 modulo M_3 as functions of those of L_1 modulo M_1 (on K_1) and L_2 modulo M_2 (on K_2). This problem can be reduced to the following: What is the point-set L_3 approximated on K_3 by the join $\bar{L}_3$ of the point-sets $\bar{L}_1$ and $\bar{L}_2$ approximating L_1 and L_2 and what is the point-set M_3 approximated by $\bar{M}_3$ which is the sum of those simplexes of $\bar{L}_3$ corresponding to a pair of simplexes of which at least one is in M_1 or M_2. For the computation as for the product and so at the same time theorem 1 can be applied on $\bar{L}_3$ modulo $\bar{M}_3 \cdot L_3$ turns out to be the set of all segments between points of L_1 and of L_2 on K_3; we can write $L_3 = L_1 \cdot L_2$. The approximation of L_3 by $\bar{L}_3$ is of a slightly different kind from that described in section 2. With the same notation $M_3 = L_1 \cdot M_2 + M_1 \cdot L_2$. The introduction of the (-1)-simplex in K_i does not have any influence on the homology characters of L_i modulo M_i provided M_i is not empty because in that case the (-1)-simplex has to be included in M_i. The repetition of Theorem 1 in its generalized form seems superfluous.

4. We base the treatment of invariant subsets on the following definition: *An invariant subset of a complex K is a subset which is transformed into itself by any transformation of K into itself or the sum of a certain number of components of such point-sets.* Combinatorial invariants of invariant subsets of K are combinatorial invariants of K. A minimal invariant subset of K is an invariant subset not containing an invariant subset different from itself. The interesting problem of finding all minimal invariant subsets of a complex is beyond the scope of the present methods of combinatorial topology.

The minimal invariant subset containing a point P of K is the component of P in the set of points of K into which P can be transformed by topological transformations of K into itself.

As any interior point of a simplex of K can be transformed into all other interior points of the same simplex, any invariant subset is the sum of the interiors of a certain number of simplexes of K. The closure of an invariant subset of K is an invariant subcomplex of K.

THEOREM 3. *There is an isotopic deformation of K into itself under which a point of any minimal invariant subset of K goes into another point of the same subset.*

To prove this we note first that there is such a deformation moving any point interior to a simplex into any other point interior to the same simplex;

second that any point of the invariant subset can be transformed into any other point of the subset. From this it follows that any point of the subset can be deformed into any point of a small neighborhood of the invariant subset and that this neighborhood can be extended to contain all interiors of simplexes belonging to the invariant subset on which, or on the boundary of which P lies. Finally it is obviously possible to pass from any point to any other point of the subset in a finite number of steps all completely contained in one of the enlarged neighborhods.

THEOREM 4. *If a simplex S^p of K^n is contained in the closure of an invariant subset M of K^n, then the point-set N, corresponding in the complement U^{n-p-1} of S^p in K^n to M, is an invariant subset of U^{n-p-1}.*

This follows from the fact that the interior points of the segments between interior points of S^p and the points of a minimal invariant subset of U^{n-p-1} all belong to the same invariant subset of K^n. It is sufficient to construct an isotopic deformation of K^n into itself moving part of such a segment into part of an arbitrary other segment, supposing that the two segments pass through the same point P of S^p and for instance through the points Q and R of U^{n-p-1}. According to theorem 3 there is an isotopic deformation of U^{n-p-1} into itself moving R into Q; this determines an isotopic deformation of the join U^{n-1} of U^{n-p-1} and the boundary of S^p into itself still moving R into Q. U^{n-1} is the neighborhood complex of P in K^n. A certain neighborhood of P, the join of P and U^{n-1}, can be formed by means of a set of copies of U^{n-1} depending on a parameter moving in an interval, however to one end-point of the interval there only corresponds the point P. We construct the isotopic deformation of K^n in such a way that under any transformation of that deformation all those copies of U^{n-1} are transformed into themselves by means of one of the transformations of the deformation of U^{n-1}. The copies near P are transformed by means of the final transformation of that deformation, the copies near the boundary of our neighborhood of P and at the same time all points outside of that neighborhood are transformed by means of the identical transformation, the copies in between those two groups are transformed by means of all the transformations of our deformation of U^{n-1} distributed continuously. The deformation of K^n is executed by moving this intervening portion continuously till it arrives at P. It is easy to see that this deformation satisfies the condition we wanted so that we have proved theorem 4.

All theorems which we are going to give with the object of finding smaller invariant subsets of a complex except theorem 10 will satisfy a certain converse of theorem 4: to any invariant point-set of a complement found by means of those theorems there will correspond an invariant point-set of the original

complex that can be found by means of the same theorems. This will be proved in section 6 and might be of use in determining invariant subsets of a given complex.

5. A large part of our method of finding invariant subsets consists in the application of the following composite statement of which we may suppress the proof:

THEOREM 5. *Invariant subsets, components, parts of certain local dimension, closures, sums, differences and intersections of invariant subsets of K^n are again invariant subsets of K^n.*

The rest of the method consists in repeated applications of generalizations of the following theorem proved in my dissertation: *

THEOREM 6. *The neighborhood complexes of corresponding points in homeomorphic complexes have the same homology characters; with the corollary: The set of points of K^n whose neighborhoods complexes have a definite set of homology characters is an invariant subset of K^n.*

The analysis of invariant subsets given by A. B. Brown † is founded on theorem 5 and a corollary of theorem 6: The subcomplex of K^n formed by all $(n-1)$-dimensional simplexes of K^n incident with exactly p $(p \neq 2)$ $n-1$ simplexes of K^n is invariant.

A first generalization of theorem 6 is the following:

THEOREM 7. *If in two homeomorphic complexes two corresponding points are interior points of p-simplexes, the complements of these p-simplexes have the same homology characters.*

Such points have as their neighborhood complexes the joins of the boundary of a p-simplex and the complements of their p-simplexes. Applying theorem 1 we find that the homology characters of those joins are obtained from the homology characters of the complements by adding p to the dimension of those characters, so that the first are invariants along with the second. From the character of the proof it is evident that this generalization does not increase our information on invariant subsets; however, it is interesting to compare theorems 7 and 10 and their proofs.

THEOREM 8. *Let M_1 and M_2 contained in M_1 be two invariant subsets of K^n and let L_1 and L_2 be the point-sets corresponding to M_1 and M_2 in the neighborhood complex of a point P of K^n, then the homology characters of L_1*

* *Loc. cit.*, p. 26.

† *American Journal of Mathematics*, Vol. 54 (1932), pp. 117-122.

modulo L_2 are invariants of K^n (and P). From this it follows: *The set of points where L_1 modulo L_2 has a definite set of homology characters is an invariant point-set of K^n.*

On the point-set K of K^n we consider as marked the following point-sets: P; the straight segments through P in K^n forming a point-set U; the sets N_1 and N_2 of those segments contained in M_1 and M_2; all point-sets corresponding to those in a complex $'K^n$ homeomorphic with K^n and marked with a prime. The intersections of U, N_1 and N_2 and a sufficiently small neighborhood of P are equal to the corresponding point-sets of $'K^n$. In a very small neighborhood of P we construct copies V and $'V$ of the neighborhood complexes of P in K^n and $'K^n$. The point-sets in those copies corresponding to the given invariant subsets will be called from now on: L_1, L_2, $'L_1$, $'L_2$. For any point of $U('U)$ except P we define a corresponding point on $V('V)$, called its projection on $V('V)$, defined as the point of intersection of the straight segment in $K^n('K^n)$ through the given point and P with $V('V)$. To any singular chain of L_1 modulo L_2 we assign its projection on $'V$ which is a chain of $'L_1$ modulo $'L_2$. As boundaries correspond to boundaries under this projection we have defined a transformation of the group of homologies of L_1 modulo L_2 into that of $'L_1$ modulo $'L_2$. To complete the proof we must show: If $'\Gamma^p$ is the projection of a cycle Γ^p of L_1 modulo L_2 on $'V$ and if $''\Gamma^p$ is the projection of $'\Gamma^p$ on V then $\Gamma^p - ''\Gamma^p \sim 0$ on L_1 modulo L_2, for then the transformation of the two groups is isomorphic. $\Gamma^p - ''\Gamma^p$ is the boundary modulo M_2 of the chain formed by the straight segments in $'K^n$ between points of Γ^p and corresponding points of $'\Gamma^p$ plus the straight segments in K^n between points of $'\Gamma^p$ and corresponding points of $''\Gamma^p$. This chain is contained in U when V and $'V$ are taken sufficiently near P; its projection on V still has $\Gamma^p - ''\Gamma^p$ as its boundary but now taken modulo L_2, so that $\Gamma^p - ''\Gamma^p \sim 0$ on L_1 modulo L_2.

Theorem 8 can be generalized to give theorem 9 in the same way that we generalized theorem 6 to give theorem 7.

THEOREM 9. *If in two homeomorphic complexes two corresponding points are interior points of p-simplexes and $L_1(L_2)$ and $'L_1('L_2)$ are the sets in the complements of these p-simplexes corresponding to the invariant subset M_1 (M_2 contained in M_1) of one complex and the corresponding set of the other complex, then the homology characters of L_1 modulo L_2 are the same as those of $'L_1$ modulo $'L_2$.*

6. *If we find by means of theorems* (5) *and* (8) *an invariant point-set L in the complement U of a p-simplex S^p of K^n then the same theorems allow*

us to construct an invariant point-set in K^n of which the dimension is $p+1$ higher and to which L corresponds in the complement of S^p.

To prove this we must follow the process by which L was found and show that corresponding steps can be taken in K^n itself with corresponding results. This is obvious as far as theorem (5) is concerned. When we apply theorem (8) using the point P of U and point-sets L_1 and L_2 in the neighborhood complex V of P in U, the corresponding point-sets for K are on the join of P, S^p and V; corresponding to P we have to take an interior point Q of the join of P and S^p, the set corresponding to $L_1(L_2)$ in the neighborhood complex of Q is the join of $L_1(L_2)$ and the boundary R^p of the $(p+1)$-simplex $P \cdot S^p$. The homology characters of $R^p \cdot L_1 \bmod R^p \cdot L_2$ are, according to Section 3, obtained from those of $L_1 \bmod L_2$ by adding $p+1$ to the dimension of every character, so that the same sets of numbers are used to determine how the points Q of K^n and the points P of U can be divided into invariant point-sets.

As a result of theorem 6 the smallest invariant subsets we can find are open manifolds in the modern sense. The invariants of tensorial character defined for manifolds by J. W. Alexander (*Proc. Nat. Acad. Sci.* 1924) can be used to find still smaller invariant subsets and a connection between the results in a complex and in the complements of its simplexes as described just now does still exist. For a subsequent paper we plan a vindication of this statement and a closer analysis of these invariants.

7. The fundamental group can be used to split a complex into still smaller invariant subsets. However, owing to the fact that deformation invariants of dimensions larger than one have not been investigated closely the range of application is much smaller than that for instance of theorem 8, and the reasoning of section 6 cannot be extended in this case.

THEOREM 10. *If the interior of a p-simplex S^p of K^n is part of a p-dimensional invariant subset of K^n, then the fundamental group of (any component of) the point-set L of the complement U of S corresponding to any other invariant point-set M of K is a combinatorial invariant of K^n (and the two invariant subsets).*

This theorem can be used to distribute the p-simplexes of a p-dimensional invariant subset over smaller invariant subsets; the lower-dimensional simplexes to be distributed afterwards by means of theorem 5.

Because S^p is in a p-dimensional invariant subset of K^n we can find an open subset of S^p of which the corresponding subset in a complex $'K^n$ homeomorphic with K^n is open in a simplex $'S^p$ of $'K^n$. For any point of the join of $S^p('S^p)$ and its complement but not on $S^p('S^p)$ itself we define a corre-

sponding point on that complement called its projection on the complement and constructed by taking the point in the complement corresponding to the join of S^p($'S^p$) and the given point. Let us represent the complements by sets on the point-set of our complex in a very small neighborhood of one of the points of the open set mentioned above so that they have a point of M in common. We chose that point in both the complements as the origin of the fundamental group of the point-sets corresponding in the complements to M. This causes the origins and so at the same time the elements of those fundamental groups to be transformed into each other under the projections defined before. Now we can repeat the substance of the proof of theorem 8, proving that L and the corresponding point-set constructed for $'K^n$ have isomorphic fundamental groups.

The restriction of our knowledge on deformation invariants explains the behavior of, for instance, the join of the boundary of a 2-simplex and a 3-dimensional manifold with a fundamental group different from the identity but whose 1-cycles are all homologous to zero. It seems very obvious that the boundary of the 2-simplex and the rest of the manifold are (minimal) invariant subsets of the join, but there does not seem to be an easy way of proving this.

8. A last theorem in this connection which we want to mention is the following:

THEOREM 11. *If in two homeomorphic n-dimensional complexes two corresponding points are both in $(n-p)$-simplexes, $p<4$, the complements of those simplexes are homeomorphic.*

The proof is so simple that we shall not go into detail. For $p=3$ the complements are divided into 2-dimensional manifolds by point-sets corresponding to $(n-1)$- and $(n-2)$-dimensional invariant subsets; the similarity of the way in which the manifolds have to be put together being decided by theorem 10 applied to those subsets ($p<3$) and the similarity of the corresponding manifolds being decided by means of theorem 9. The sign of the relation between any orientable manifold and a 1-dimensional invariant part of its boundary is not changed in passing to a homeomorphic complex because the sign of the relation between the chain formed by the manifold modulo all the rest of the boundary and the cycle formed by the chosen part of the boundary modulo its boundary is not changed in passing to a homeomorphic complex.

THE JOHNS HOPKINS UNIVERSITY.

SOME ADDITIONS TO THE THEORY OF COMBINATORS.*

By H. B. CURRY.†

The present paper contains two amplifications of my dissertation *Grundlagen der kombinatorischen Logik*.‡ These are: (1) a revision of the definition of normal form (II D 1, Festsetzung 2) so that to every normal sequence there will correspond a unique normal combinator; (2) a generalization of the principal theorems on equality of combinators (II E 3, Sätze 4 and 5), so that the restriction that the combinators concerned be "eigentlich" is removed. These two parts of this paper are independent of one another, although both presuppose the thesis just mentioned.

Reference will be made to the dissertation by means of the subdivisions in that paper itself. Thus the above II D 1 means Kapitel II, Abschnitt D, § 1.

1. *Revised Definition of Normal Form.*

Convention 1. A regular combinator § whose terms are all of the form C_m, or $B_m I$) shall be called a *permutator*. (By II D 1, Festsetzung 4, permutators were denoted by $\mathfrak{C}$).

Convention 2. A permutator shall be called *normal*, or in the normal form, when it is either I or of the form

$$\mathfrak{C}_n \cdot \mathfrak{C}_{n-1} \cdot \mathfrak{C}_{n-2} \cdot \cdot \cdot \mathfrak{C}_2 \cdot \mathfrak{C}_1,$$

where each $\mathfrak{C}_k$ either is of the form

$$C_{a_k} \cdot C_{a_{k+1}} \cdot C_{a_{k+2}} \cdot \cdot \cdot C_{k-1}, \qquad (a_k < k),$$

or else, for $k \neq n$, $\mathfrak{C}_k \equiv I$.

THEOREM 1. *For every permutator* $\mathfrak{C}$ *there exists one and only one normal permutator* $\mathfrak{C}'$ *such that*

$$\vdash \mathfrak{C} = \mathfrak{C}'. \tag{1}$$

Proof. If such a $\mathfrak{C}'$ exists it corresponds to a permutation (II C 4, Festsetzung 3) of order n (II D 6, Festsetzung 1). n must then also be the

* Presented to the American Mathematical Society, December 28, 1931.

† National Research Fellow.

‡ *American Journal of Mathematics*, Vol. 52 (1930), pp. 509-536, pp. 789-834.

§ II D 1, Festsetzung 1.

order of the permutation to which $\mathfrak{C}$ corresponds, (II D 6, Satz 3); hence n is uniquely determined.*

If U is any permutator let us denote by $U(k)$ the index of the variable which occupies the k-th place in the permutation to which U corresponds. Thus if $b_k = \mathfrak{C}(k)$ we have, if the x's be regarded as entities,

$$\vdash \mathfrak{C}x_0x_1x_2 \cdots x_n = x_0x_{b_1}x_{b_2} \cdots x_{b_n}.$$

I shall show that the numbers a_k determined as in Convention 2 by $\mathfrak{C}'$ are uniquely determined by the b_k, with the understanding that when $a_k = k$, then $\mathfrak{C}_k \equiv I$.

Suppose then a $\mathfrak{C}'$ is given, with the a_k defined in terms of $\mathfrak{C}'$ as in Convention 2 with the understanding just mentioned. Then

$$\begin{aligned} &\mathfrak{C}_k(j) = j, \quad \text{if} \quad j > k; \\ &\mathfrak{C}_k(k) = a_k. \end{aligned} \tag{2}$$

Now it is easily verified that for permutators U and V

$$(U \cdot V)(k) = U(V(k));$$

hence

$$(\mathfrak{C}_k \cdot \mathfrak{C}_{k-1} \cdots \mathfrak{C}_1)(k) = a_k. \tag{3}$$

Consequently, if we define temporarily

$$D_k \equiv \mathfrak{C}_n \cdot \mathfrak{C}_{n-1} \cdots \mathfrak{C}_{k+1}, \qquad D_n \equiv I, \tag{4}$$

so that

$$\vdash \mathfrak{C}' = D_k \cdot (\mathfrak{C}_k \cdot \mathfrak{C}_{k-1} \cdots \mathfrak{C}_1),$$

then, by (2), (3),

$$\mathfrak{C}'(k) = D_k(a_k) = D_{k+1}(k). \tag{5}$$

If $\mathfrak{C}'$ is to satisfy the conditions, $\mathfrak{C}'(k)$ must equal b_k, hence

$$b_k = D_k(a_k),$$

or, since permutators have unique universes,

$$a_k = D_k^{-1}(b_k). \tag{6}$$

The equations (6) determine the a's uniquely. In fact I shall prove that for any $r < n$ there exists a unique set of numbers $a_n, a_{n-1}, \cdots, a_{n-r}$ such that: (1) $a_n < n$, and for $k \leqq r$, $a_{n-k} \leqq n - k$; (2) if we define $\mathfrak{C}_{n-k}$ as in Convention 2 for $a_{n-k} < n - k$, and as I for $a_{n-k} = n - k$, then the last $r + 1$

* These statements are true even when $\mathfrak{C}' \equiv I$, $n = 0$.

equations (6) are satisfied, D_k being defined by (4). This assertion is true for $r = 0$; for the last equation (6) is

$$a_n = b_n,$$

which determines a_n uniquely; moreover $a_n < n$ since $\mathfrak{C}$ corresponds to a sequence of order n. Suppose now the assertion is true for a given r. Then the last $r+1$ equations (6) determine $a_n, a_{n-1}, \cdots, a_{n-r}$ uniquely and subject to the inequalities stated. The next preceding equation, viz.,

$$a_{n-r-1} = D^{-1}{}_{n-r-1}(b_{n-r-1}),$$

determines a_{n-r-1} uniquely, since $D^{-1}{}_{n-r-1}$ contains no number not already determined, and $a_{n-r-1} \leqq n$. Suppose $a_{n-r-1} = j > n-r-1$. Then

$$\begin{aligned} D_{n-r-1}(j) &= (D_j \cdot \mathfrak{C}_j \cdot \mathfrak{C}_{j-1} \cdot \cdots \mathfrak{C}_{n-r})(j) && \text{(by def. of } D_{n-r-1}), \\ &= (D_j \cdot \mathfrak{C}_j)(j) && \text{(by (2))}, \\ &= D_j(a_j) = b_j && \text{(by (5))}. \end{aligned}$$

But,

$$(7) \qquad D_{n-r-1}(j) = b_{n-r-1} \qquad \text{(by (6))}.$$

$\therefore b_j = b_{n-r-1}$, which is impossible if $j > n-r-1$ because the b's are all distinct. Therefore $a_{n-r-1} \leqq n-r-1$, and the above assertion, if true for r, is also true for $r+1$. By induction it is therefore true for every $r < n$; hence in particular for $r = n-1$. The equations (6) have therefore one and only one set of solutions $a_1, a_2, \cdots, a_n$.

Now let the $a_1, a_2, \cdots, a_n$ be determined as in the last paragraph, and let $\mathfrak{C}'$, D_k, be defined in terms of them. Then by (5), which is still true, $\mathfrak{C}'$ satisfies the conditions of the theorem, and is by what has just been proved the only normal permutator which does so.

Convention 3. A regular combinator shall be said to be *normal*, or in the normal form, when (except for parentheses and the omission of identical terms) it is in the form

$$\mathfrak{A} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B} *$$

(1) $\mathfrak{A}$, $\mathfrak{W}$, $\mathfrak{B}$, are in the normal forms already defined (II D 1, Festsetzung 2);

(2) $\mathfrak{C}$ is in the form of Convention 2;

(3) $\mathfrak{C}$ corresponds to the permutation π determined as follows: let η be

* Any of the components $\mathfrak{A}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$, may reduce to the identical combinator, I, and hence be omitted.

the sequence (Folge) to which the given combinator corresponds; determine from η a γ and an ω as in II C 5, Satz 1; from ω a κ and a μ as in II C 4, Satz 1; and from μ an α satisfying the conditions for the first factor in II C 4, Satz 3; then of those permutations which satisfy the conditions for the second factor of the last theorem choose for π that one which does not permute among themselves any of the letters $x_{c_1}, x_{c_2}, \cdots x_{c_p}$, which occupy places held by the same letters in α.*

THEOREM 2. *To every normal sequence there corresponds one and only one normal combinator.*

Proof. Let η be the given sequence, and let γ, ω, κ, μ, α, π be determined by the process outlined in Convention 3. Then

$$\eta = \kappa \cdot \alpha \cdot \pi \cdot \gamma.$$

Let $\mathfrak{K}$, $\mathfrak{W}$, $\mathfrak{C}$, $\mathfrak{B}$ be the normal combinators which correspond to κ, α, π, and γ, respectively (II C 4, Sätze 2 and 4, Theorem 1 above, and II C 3, Satz 3). Then $(\mathfrak{K} \cdot \mathfrak{W} \cdot \mathfrak{C} \cdot \mathfrak{B})$ is normal and corresponds to η.

Suppose there were another such combinator; let it be $(\mathfrak{K}' \cdot \mathfrak{W}' \cdot \mathfrak{C}' \cdot \mathfrak{B}')$. By II C 5, Satz 2, these combinators can differ only in their $\mathfrak{C}$'s. By the third requirement in Convention 3, both $\mathfrak{C}$ and $\mathfrak{C}'$ must correspond to π as described in that requirement. By the theorems there mentioned, ω, μ, α, and therefore π, are uniquely determined by η.† Hence $\mathfrak{C}$ and $\mathfrak{C}'$ must both correspond to the same permutation, and therefore, by Theorem 1, are identical.

THEOREM 3. *If $\mathfrak{R}$ is a regular combinator, then there exists one and only one normal combinator $\mathfrak{R}'$ such that*

$$\vdash \mathfrak{R} = \mathfrak{R}'.$$

Proof. Follows at once by II D 6, Sätze 2, 3, 4, and Theorem 2.

2. *General Theorem on Equality of Combinators.*

Convention 4. If $\mathfrak{X}$ and $\mathfrak{Y}$ are expressions involving the letters $x_1, x_2, \cdots, x_m$; then

$$\vdash_x \mathfrak{X} = \mathfrak{Y}$$

---

* This π is the same as the β constructed as an example in my paper "An Analysis of Logical Substitution," *American Journal of Mathematics*, Vol. 51 (1929), pp. 363-384, Lemma 4; it being understood that the γ of that paper is the same as the μ mentioned here. The other permutations which satisfy the conditions for the second factor in II C 4, Satz 3 are the β's discussed in Lemma 5 of the earlier paper, from which it appears that π is uniquely described.

† See footnote above.

shall mean that if the $x_1, x_2, \cdots x_m$ be treated as variables,—i. e. as denoting indeterminate entities,—then the equality of $\mathfrak{X}$ and $\mathfrak{Y}$ can be proved. Similarly if $\mathfrak{X}$ and $\mathfrak{Y}$ also contain the letters $y_1, y_2, \cdots, y_n$, $\vdash_{x,y} \mathfrak{X} = \mathfrak{Y}$ shall mean that the equality can be proved if the x's and y's be treated as variables. (These conventions apply even when the x's and y's actually denote entities. Cf. II E 3, Festsetzung 3).

Convention 5. A formula, such as $\vdash X = Y$, shall be said to *follow combinatorially* when it can be derived, as in II, from the combinatory axioms, the combinatory rules *B*, *C*, *W*, *K*, and the properties of equality proved in *I D*. This may be symbolized thus

$$\vdash X = Y \qquad \text{(combinatorially)}.$$

The same shall apply to formulas of type considered in Convention 4.

THEOREM 4. *If* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *are combinations of variables* $x_1, x_2, \cdots, x_n$, *and entities* $u_1, u_2, \cdots, u_m$,* *and* X *and* Y *are combinators such that,—*

(i) *the expressions* $(XIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ *and* $(YIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ *reduce formally to* $\mathfrak{X}$ *and* $\mathfrak{Y}$ *respectively;*

(ii) *it follows combinatorially that*

$$\vdash_x \mathfrak{X} = \mathfrak{Y}; \tag{1}$$

then it follows combinatorially that

$$\vdash XIu_1u_2 \cdots u_m = YIu_1u_2 \cdots u_m. \tag{2}$$

Proof. Suppose first that $\mathfrak{X}$ differs from $\mathfrak{Y}$ only in that a certain expression $\mathfrak{A}$ in $\mathfrak{X}$ is replaced in $\mathfrak{Y}$ by $\mathfrak{B}$, where

$$\vdash_x \mathfrak{A} = \mathfrak{B} \qquad \text{(combinatorially)}. \tag{3}$$

Then there must exist a regular combinator T such that the expressions $(TI\mathfrak{A}u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ and $(TI\mathfrak{B}u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively.† We shall consider two cases.

Case 1. Suppose $\mathfrak{A}$ and $\mathfrak{B}$ are entities, and that the formula (3) is established without the use of variables. Let U and V be combinators such

---

* It is of course not required that all these symbols appear in both $\mathfrak{X}$ and $\mathfrak{Y}$.

† For this reduction $\mathfrak{X}$ is to be regarded as a combination of the symbols $\mathfrak{A}, u_1, u_2, \cdots, u_m, x_1, x_2, \cdots, x_n$, and $\mathfrak{A}$ appears only once in $\mathfrak{X}$. Similarly for the relation between $\mathfrak{Y}$ and $\mathfrak{B}$. Superfluous u's can be cancelled by factors K_j appearing in T.

that $UIu_1u_2 \cdots u_m$ and $VIu_1u_2 \cdots u_m$ formally reduce respectively to $\mathfrak{A}$ and $\mathfrak{B}$, and let S be a regular combinator corresponding to the sequence

$$x_0x_2(x_1x_2x_3x_4 \cdots x_{m+2})x_3x_4 \cdots x_{m+2}x_{m+3} \cdots .$$

Let $X' \equiv STU$, $Y' \equiv STV$. Then the expressions $(X'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ and $(Y'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n)$ reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively.* Therefore, by II E 3, Satz 4, $\vdash XI = X'I$ and $\vdash YI = Y'I$, because, for example, if the u's be regarded as variables the two expressions

$$XIu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n \quad \text{and} \quad X'Iu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$$

both reduce to the same sequence of variables and hence XI and $X'I$ correspond to the same sequence of variables. On the other hand

$$\vdash X'Iu_1u_2 \cdots u_m = TI\mathfrak{A}u_1u_2 \cdots u_m,$$
$$\vdash Y'Iu_1u_2 \cdots u_m = TI\mathfrak{B}u_1u_2 \cdots u_m.$$

Since the two left-hand sides reduce formally to their respective right-hand sides. The two right-hand sides are equal by (3). Hence, finally,

$$\begin{aligned} \vdash XIu_1u_2 \cdots u_m &= X'Iu_1u_2 \cdots u_m \\ &= Y'Iu_1u_2 \cdots u_m \\ &= YIu_1u_2 \cdots u_m. \end{aligned}$$

which proves the theorem for this case.

Case 2. Suppose $\mathfrak{A}$ differs from $\mathfrak{B}$ in that $\mathfrak{A}$ reduces to $\mathfrak{B}$ by application of a single one of the combinatory rules, B, C, W, K. (There is no loss of generality in supposing that $\mathfrak{A}$ reduces to $\mathfrak{B}$). Let U be the combinator concerned in this reduction; it is then one of the u_i. Let $\mathfrak{X}'$ be the expression obtained from $\mathfrak{X}$ by replacing U as it appears at the beginning of $\mathfrak{A}$ by a new variable y, but otherwise leaving $\mathfrak{X}$ unchanged.† Let X' be the regular combinator such that $X'Iyu_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$ reduces formally to $\mathfrak{X}'$. Let V be the regular combinator such that

$$\vdash_{x,u} Vx_0u_1u_2 \cdots u_m = x_0Uu_1u_2 \cdots u_m,$$

(such a V exists since U is simply one of the u_i). Then the expression

* The first one reduces formally to $TI(UIu_1u_2 \cdots u_m)u_1u_2 \cdots u_m\, x_1x_2 \cdots x_n$, thence to $\mathfrak{X}$ with $\mathfrak{A}$ replaced by $UIu_1u_2 \cdots u_m$. In this the only symbol not considered formally is U, so that the reduction proceeds formally to $\mathfrak{X}$. Similarly for the second one.

† If U appears in $\mathfrak{X}$ in more than one place, this replacement is to be made only at the beginning of $\mathfrak{A}$.

$(V \cdot X')Iu_1u_2 \cdot \cdot \cdot u_m x_1x_2 \cdot \cdot \cdot x_n$ reduces formally to $\mathfrak{X}$ (the u_i, x_j all being regarded as variables). Hence by II E 3, Satz 4, (cf. above)

$$\vdash (V \cdot X')I = XI. \tag{4}$$

On the other hand if we regard U as a combinator and $u_1, \cdot \cdot \cdot, u_m, x_1, \cdot \cdot \cdot x_n$, as variables, the expression $X'IUu_1u_2 \cdot \cdot \cdot u_p x_1x_2 \cdot \cdot \cdot x_n$ reduces to $\mathfrak{Y}$. Hence, (II E 3, Satz 4)

$$\vdash X'IU = YI. \tag{5}$$

Therefore

$$\begin{aligned} \vdash XIu_1u_2 \cdot \cdot \cdot u_m &= (V \cdot X')Iu_1u_2 \cdot \cdot \cdot u_m && \text{(By (4)),} \\ &= V(X'I)u_1u_2 \cdot \cdot \cdot u_m && \text{(II B 4, Satz 1),} \\ &= X'IUu_1u_2 \cdot \cdot \cdot u_m && \text{(Def. of } V\text{),} \\ &= YIu_1u_2 \cdot \cdot \cdot u_m && \text{(By (5)).} \end{aligned}$$

The theorem is therefore proved in case the transition from $\mathfrak{X}$ to $\mathfrak{Y}$ is effected by means of a single formula of type (3). For if (3) be an axiom or a previously proved formula we have the Case 1; if (3) be a special case of a rule, Case 2.

In the most general case there must exist a series of expressions $\mathfrak{X}_1, \mathfrak{X}_2, \cdot \cdot, \mathfrak{X}_p$, where $\mathfrak{X}_1 \equiv \mathfrak{X}$, $\mathfrak{X}_p \equiv \mathfrak{Y}$, such that $\mathfrak{X}_{i+1}$ is derived from $\mathfrak{X}_i$ by a single formula (3) in the manner just considered. If X_i be a combinator related to $\mathfrak{X}_i$ in the same manner as X is related to $\mathfrak{X}$ and Y to $\mathfrak{Y}$,* then by the preceding proof,

$$\vdash X_iu_1u_2 \cdot \cdot \cdot u_m = X_{i+1}u_1u_2 \cdot \cdot \cdot u_m.$$

Therefore, since equality is transitive,

$$\vdash Xu_1u_2 \cdot \cdot \cdot u_m = Yu_1u_2 \cdot \cdot \cdot u_m. \qquad \text{q. e. d.}\dagger$$

---

* Such combinations of course exist; in fact we can find an X_i of the form $\mathfrak{N}_iI$, where $\mathfrak{N}_i$ is normal.

† In case the $\mathfrak{X}_i$ contain letters appearing in neither $\mathfrak{X}$ nor $\mathfrak{Y}$ the following modification of the proof is required. Let $v_1, v_2, \cdot \cdot \cdot, v_p$, $y_1, y_2, \cdot \cdot \cdot, y_q$ be the extra letters so appearing, the v's representing entities and the y's being variables. Let

$$V_r = K_r \cdot K_{r-1} \cdot \cdot \cdot K_2 \cdot K_1, \qquad r = p + q.$$

Then by the method in the text we can establish the formula

$$\vdash (B_mV_r \cdot X)Iu_1u_2 \cdot \cdot \cdot u_m v_1v_2 \cdot \cdot \cdot v_p = (B_mV_r \cdot Y)Iu_1u_2 \cdot \cdot \cdot u_m v_1v_2 \cdot \cdot \cdot v_p,$$

for if we insert first the y's and then the x's after the two sides the resulting expressions reduce formally to $\mathfrak{X}$ and $\mathfrak{Y}$ respectively. Hence if $w_1, w_2, \cdot \cdot \cdot, w_q$ are any entities (say $w_i \equiv I$), we have

$$\begin{aligned} \vdash (B_mV_r \cdot X)Iu_1u_2 \cdot \cdot \cdot u_m v_1 \cdot \cdot \cdot v_p w_1 \cdot \cdot \cdot w_q \\ = (B_mV_r \cdot Y)Iu_1u_2 \cdot \cdot \cdot u_m v_1 \cdot \cdot \cdot v_p w_1 \cdot \cdot \cdot w_q, \end{aligned}$$

whence, by reduction of both sides, we have (2).

Remark. That the above theorem is more general than those of II E 3, is shown by the following example, which may be proved by its aid,

$$\vdash C(W \cdot BB)I = WCI.$$

Neither of the combinators is "eigentlich"; in fact if x_0x_1 be written after each they both reduce to x_0Ix_1. On the other hand if I is to be regarded formally the right-hand side reduces to the expression just written, while the reduction of the left-hand side cannot get any further than $x_0I(Ix_1)$.

THEOREM 5. *If X any Y are entities such that it follows combinatorially that*

$$\vdash_x Xx_1x_2 \cdots x_n = Yx_1x_2 \cdots x_n;$$

then it follows combinatorially that

$$\vdash X = Y.$$

Proof. Let X' and Y' be combinators, and $u_1, u_2, \cdots u_m$ entities such that $X'Iu_1u_2 \cdots u_m$ and $Y'Iu_1u_2 \cdots u_m$ reduce formally to X and Y respectively. Then

$$\begin{aligned} \vdash X &= X'Iu_1u_2 \cdots u_m && \text{(by const.),} \\ &= Y'Iu_1u_2 \cdots u_m && \text{(Theorem 4),} \\ &= Y \qquad \text{q. e. d.} && \text{(by const.).} \end{aligned}$$

UNIVERSITY OF CHICAGO.

ON THE OVERCONVERGENCE OF SEQUENCES OF RATIONAL FUNCTIONS.*

By J. L. WALSH.

If a sequence of rational functions of the complex variable z converges to a given function $f(z)$ on a given point set C with a certain degree of approximation, then that sequence may necessarily converge to the given function or its analytic extension on a larger point set containing C in its interior. For instance, it was shown by S. Bernstein that convergence of the sequence of polynomials $p_n(z)$ of respective degrees $n = 1, 2, \cdots$ in such a way that we have for some $R > 1$

$$| f(z) - p_n(z) | \leqq M/R^n, \qquad (-1 \leqq z \leqq 1),$$

implies uniform convergence of the sequence $p_n(z)$ in some ellipse whose foci are the points 1 and -1. Various generalizations of this result have been studied by the present writer, where the degree of approximation of polynomials is given in a certain way not on the point set $-1 \leqq z \leqq 1$, but on a more general point set,† and also where the approximating functions are rational functions more general than polynomials and the point set is a region or rectifiable curve.‡ It is the object of the present paper to set forth a result which includes all of those mentioned, and which indeed seems to be the most general result of its type. In the previous study, we considered *primarily* approximation in a circular region by rational functions whose poles lie in another circular region; in the present paper more general point sets instead of circular regions are contemplated in these two rôles. We shall first prove an interesting special case, and then indicate the modifications for the more general theorem.

I.

Let C be an arbitrary closed Jordan region of the extended § z-plane and let D be an arbitrary closed point set having no point in common with C.

---

* Presented to the Society, March 25, 1932.

† *Münchner Berichte* (1926), pp. 223-229.

‡ *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 838-847 and Vol. 34 (1932), pp. 22-74.

§ That is to say, we adjoin the point at infinity to the usual (finite) z-plane; consequently the notions *closure* (of a point set), *Jordan curve*, and *Jordan region* are most conveniently interpreted on the sphere (stereographic projection of the plane) instead of in the plane itself.

Let $C_R(z')$ denote the curve in the z-plane which is the image of the circle $|w| = R > 1$ when the complement of C is mapped onto $|w| > 1$ in such a way that the point $z = z'$ of D corresponds to $w = \infty$. The closed regions containing C bounded by the various curves $C_R(z')$ have obviously the region C in common when z' varies over D, and conceivably have a still larger point set in common. We denote by C_R this common point set whatever it may be. It will appear later that the point set common to the regions considered is itself closed, and it will also appear that C actually lies interior to C_R.

LEMMA I. *If $r_n(z)$ is a rational function of z of degree n * whose poles lie in D and if we have*

$$|r_n(z)| \leqq M, \quad \text{for } z \text{ on } C, \tag{1}$$

then we have also

$$|r_n(z)| \leqq MR^n, \quad \text{for } z \text{ on } C_R. \tag{2}$$

Let $z_1, z_2, \cdots, z_n$ be the poles of $r_n(z)$; these lie in D and need not be distinct. The function $r_n(z)$ need not have n poles; if there are fewer than n poles, only an obvious modification is to be made in our reasoning. Let $w = \phi_i(z)$ be a function which maps the exterior of C onto the exterior of $\gamma: |w| = 1$ so that the point $z = z_i$ corresponds to $w = \infty$. The function $\phi_i(z)$ is analytic exterior to C and (if suitably defined on C itself) continuous in the corresponding closed region except for a pole of the first order at $z = z_i$. The modulus of $\phi_i(z)$ on C is unity. The function

$$r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_n(z)$$

is analytic exterior to C, continuous in the corresponding closed region, and for z on C we have by (1)

$$|r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_n(z)| \leqq M. \tag{3}$$

Since inequality (3) holds for z on C it holds also for z exterior to C. Throughout C_R exterior to C we have

$$|\phi_i(z)| \leqq R,$$

so that (2) now follows from (3) for z in C_R exterior to C. But C_R contains C in its interior, and $r_n(z)$ has no poles in C_R, so (2) is valid for z anywhere in C_R.

---

* That is to say, $r_n(z)$ can be written in the form

$$(a_0z^n + a_1z^{n-1} + \cdots + a_n)/(b_0z^n + b_1z^{n-1} + \cdots + b_n),$$

where the denominator does not vanish identically.

Lemma I is analogous to, and a generalization of, a lemma used by Bernstein in the proof of his result already mentioned. The present method is closely related to that used by M. Riesz * in his proof of Bernstein's lemma.

Lemma I can obviously be extended to apply to point sets C more general than Jordan regions; we shall discuss this fact in more detail later. Let us return to the original case considered and use our original notation.

LEMMA II. *If C is a Jordan region and if the point set D (having no point in common with C) is a Jordan region bounded by a curve which can be denoted by $C_\rho(\zeta)$ [where ζ is some point of D], then the point set C_R is the Jordan region containing C bounded by the curve $C_\nu(\zeta)$, where*

$$\nu = (R\rho + 1)/(R + \rho).$$

Let $w = \phi(z)$ be a function which maps the exterior of C onto the exterior of $\gamma: |w| = 1$ so that the point $z = \zeta$ corresponds to $w = \infty$. Then the function

$$w' = (1 - \bar{w}_1 w)/(w - w_1) = [1 - \overline{\phi(z_1)}\phi(z)]/[\phi(z) - \phi(z_1)]$$

maps the exterior of C onto the exterior of $\gamma': |w'| = 1$ so that $z = z_1$ corresponds to $w' = \infty$. The restrictions $|w_1| \geqq \rho$, $|\phi(z_1)| \geqq \rho$, z_1 in the closed region not containing C bounded by $C_\rho(\zeta)$, z_1 in D, are all equivalent. All points z not in C_R can be expressed by the conditions

$$\left| \frac{1 - \overline{\phi(z_1)}\phi(z)}{\phi(z) - \phi(z_1)} \right| \geqq R, \qquad |\phi(z_1)| \geqq \rho, \tag{4}$$

or by the equivalent conditions

$$\left| \frac{1 - \bar{w}_1 w}{w - w_1} \right| \geqq R, \qquad |w_1| \geqq \rho, \tag{5}$$

where $w = \phi(z)$, $w_1 = \phi(z_1)$.

Conditions (5), considered as restrictions on w, are not difficult to transform. We substitute

$$w' = (1 - \bar{w}_1 w)/(w - w_1),$$

and write (5) in the form

$$w = (w_1 w' + 1)/(w' + \bar{w}_1), \qquad |w'| \geqq R > 1, \qquad |w_1| \geqq \rho > 1. \tag{6}$$

Here w' and w_1 are arbitrary, subject to the conditions indicated, and we seek the locus of w. The transformation in (6) transforms the unit circle

* *Acta Mathematica*, Vol. 40 (1916), pp. 337-347.

$| w' | = 1$ and its exterior into $| w | = 1$ and its exterior, so we obviously have $| w | > 1$. Let us set $| w' | \geqq R_1 \geqq R$, $| w_1 | = \rho_1 \geqq \rho$, so we have by (6)

$$w\bar{w} = (R_1^2\rho_1^2 + w_1w' + \bar{w}_1\bar{w}' + 1)/(R_1^2 + w_1w' + \bar{w}_1\bar{w}' + \rho_1^2).$$

This last expression can be written as

$$\frac{R_1^2\rho_1^2 + 2x + 1}{R_1^2 + 2x + \rho_1^2}, \tag{7}$$

a fraction whose numerator is greater than its denominator, which is positive. We are considering the fraction for $-\rho_1R_1 \leqq x \leqq \rho_1R_1$. The denominator vanishes for $x = -(R_1^2 + \rho_1^2)/2 \leqq -R\rho$, so, since the function (7) is represented by an equilateral hyperbola whose asymptotes are the axes, its minimum occurs for $x = \rho_1R_1$, and this minimum value is

$$[(R_1\rho_1 + 1)/(R_1 + \rho_1)]^2.$$

The relation

$$(R_1\rho_1 + 1)/(R_1 + \rho_1) \geqq (R\rho + 1)/(R + \rho)$$

is a consequence of the conditions $R_1 \geqq R$, $\rho_1 \geqq \rho$, so we have shown that *the complement of C_R lies in the Jordan region not containing C bounded by the curve $C_\nu(\zeta)$, $\nu = (R\rho + 1)/(R + \rho)$.*

Reciprocally, let z be given in the Jordan region not containing C bounded by this curve $C_\nu(\zeta)$; we shall prove that z belongs to the complement of C_R. When this situation is transformed onto the w-plane, it reduces to the following. Given w, in absolute value not less than $(R\rho + 1)/(R + \rho)$; to find w' and w_1 such that we have (6) satisfied.

If $| w | \geqq R$, it suffices to set $w' = w$, $w_1 = \infty$. If we have

$$R > | w | \geqq (R\rho + 1)/(R + \rho),$$

we define the quantity ρ_1 by the equation

$$| w | = (R\rho_1 + 1)/(R + \rho_1)$$

so we have $\rho_1 \geqq \rho$. Then (6) is satisfied if we merely set

$$w_1 = \rho_1 w(R + \rho_1)/(R\rho_1 + 1), \qquad w' = R(R\rho_1 + 1)/w(R + \rho_1);$$

we have $| w_1 | = \rho_1$, $| w' | = R$, and Lemma II is completely proved.

It follows from Lemma II that the complement of C_R never has a point in common with C, no matter how D (having no point in common with C) may be chosen. For such an arbitrary D lies in some Jordan region D' bounded by a Jordan curve $C_\rho(\zeta)$. Enlargement of the original D to the new D' merely adds points to the complement of C_R or leaves that complement

unchanged, and the complement of the C_R corresponding to D' has no point in common with C. Thus C itself always lies interior to C_R.

It is also true under our present hypothesis, namely that C is a Jordan region and D an arbitrary closed set having no point in common with C, that the point set C_R is closed. In fact, C_R is the point set common to a finite or infinite number of closed regions (each bounded by a curve $C_R(z')$ and containing C) and therefore closed.

It is not true under our present hypothesis that C_R is necessarily connected, as we proceed to illustrate by an example. Let us choose C as the unit circle, and D as the point set 2, 2ω, $2\omega^2$, where ω is an imaginary cube root of unity. The curve $C_R(z')$ is a circle of the coaxial family determined by z' (considered as a null circle) and C. For large values of R, the three circles $C_R(2)$, $C_R(2\omega)$, $C_R(2\omega^2)$ are small circles which do not intersect each other and which contain the points 2, 2ω, $2\omega^2$ respectively in their interiors. As R decreases, those circles vary as follows: they increase in size, intersect each other, become straight lines, become circles not containing the respective points 2, 2ω, $2\omega^2$ but containing C, decreases in size, and finally approach C. At the stage where these circles $C_R(z')$ are still circles containing the respective points 2, 2ω, $2\omega^2$ but intersecting each other, the point at infinity is exterior to the circles $C_R(z')$ and hence a point of C_R, yet cannot be connected with C by a broken line consisting wholly of point of C_R.

By means of Lemma I we shall be able to prove

THEOREM I. *Let $r_n(z)$ be a sequence of rational functions of respective degrees n such that we have for z in an arbitrary closed Jordan region C*

$$|f(z) - r_n(z)| \leqq M/\rho^n, \qquad \rho > 1. \tag{8}$$

If the poles of the functions $r_n(z)$ lie in the closed set D, where D has no point in common with C, then the sequence $r_n(z)$ converges uniformly for z in C_R, where $R < \rho^{1/2}$.

If (8) is valid for z in C and if the function $r_{n+1}(z) - r_n(z)$ is of degree $n+1$, with its poles in D, then the sequence $r_n(z)$ converges uniformly for z in C_R, where $R < \rho$.

If (8) is valid for z in C and if there exists a sequence $r'_n(z)$ of functions such that $r'_n(z) - r_n(z)$ is a rational function of degree n with its poles in D, if we have

$$|f(z) - r'_n(z)| \leqq M/\rho^n, \qquad z \text{ in } C, \tag{9}$$

and if the sequence $r'_n(z)$ converges uniformly for z in C_R, where $R < \rho$ (R may or may not depend on the sequence $r'_n(z)$), then the sequence $r_n(z)$ also converges uniformly for z in C_R.

We prove the various parts of Theorem I in order. From the two inequalities

$$| f(z) - r_n(z) | \leqq M/\rho^n, \qquad | f(z) - r_{n+1}(z) | \leqq M/\rho^{n+1},$$

satisfied for z on C, we have

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right),$$

also for z on C. The rational function $r_{n+1}(z) - r_n(z)$ is of degree $2n+1$, so Lemma I yields for z on C_R,

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right) R^{2n+1} = MR\left(1 + \frac{1}{\rho}\right)\left(\frac{R^2}{\rho}\right)^n.$$

Hence the sequence $r_n(z)$ converges as we have asserted.

It follows also from the proof just given that if $r_{n+1}(z) - r_n(z)$ is of degree $n+1$, with its poles in D, we have for z on C_R

$$| r_{n+1}(z) - r_n(z) | \leqq \frac{M}{\rho^n}\left(1 + \frac{1}{\rho}\right) R^{n+1},$$

so the sequence $r_n(z)$ converges uniformly for z on C_R, $R < \rho$. The assumption that $r_{n+1}(z) - r_n(z)$ is a rational function of degree $n+1$ is satisfied for instance by the sequence of functions corresponding to a series of the type*

$$a_0 + a_1 \frac{z-\beta_1}{z-\alpha_1} + a_2 \frac{(z-\beta_1)(z-\beta_2)}{(z-\alpha_1)(z-\alpha_2)} + a_3 \frac{(z-\beta_1)(z-\beta_2)(z-\beta_3)}{(z-\alpha_1)(z-\alpha_2)(z-\alpha_3)} + \cdots$$

If we have both (8) and (9) valid, we obtain similarly for z on C,

$$| r_n(z) - r'_n(z) | \leqq 2M/\rho^n,$$

and Lemma I gives the inequality

$$| r_n(z) - r'_n(z) | \leqq (2M/\rho^n) R^n$$

for z on C_R. Thus the sequence $r_n(z) - r'_n(z)$ converges uniformly for z on C_R, $R < \rho$. The sequence $r'_n(z)$ is known to converge uniformly for z on C_R, so it follows that the sequence $r_n(z)$ converges also in this manner, and the proof of Theorem I is complete. If (9) is used, we do not need to know that $r_n(z)$ is rational or of degree n.

In each of the parts of Theorem I it is naturally true that the sequence $r_n(z)$ converges to the function $f(z)$ (or its analytic extension) for z in C_R if C_R is connected and otherwise for z in that connected part of C_R containing C; for by (8) itself the limit of the sequence is $f(z)$ on C, and the limit of the sequence is analytic for z on C_R.

---

* Compare Angelescu, *Bulletin de l'Académie Roumaine*, Vol. 9 (1925), pp. 164-168; Walsh, *Proceedings of the National Academy of Sciences*, Vol. 18 (1932), pp. 165-171.

II.

It is clear that the reasoning used in proving Lemmas I and II can be extended to include point sets C much more general than Jordan regions; the reasoning holds with only minor changes if C is any region or other closed point set whose complement is connected. But if the complement of C is not connected, our previous discussion requires to be supplemented, and we proceed to indicate how that can be accomplished.

Let C and D be arbitrary closed point sets with no common point, such that the Dirichlet problem (for arbitrary boundary values) has a solution for every region $S(z')$ defined as the totality of points which can be joined to a given point z' of D by broken lines lines not meeting C. Such a totality of points actually forms an open region and its boundary consists entirely of points of C, as the reader will easily prove.* As before, let $C_R(z')$ denote the curve or curves in the z-plane which are the image of the circle $| w | = R < 1$ when the region $S(z')$ is mapped onto $| w | > 1$ in such a way that the point $z = z'$ corresponds to $w = \infty$. The mapping need not be smooth and the mapping function $w = \phi(z)$ is not uniquely defined, but the mapping function exists by the hypothesis on the region $S(z')$, and the point set $C_R(z')$ is uniquely defined. Let $S_R(z')$ denote the complement in the z-plane of the set $| \phi(z) | > R > 1$, so that $S_R(z')$ is closed and contains C, and is bounded by $C_R(z')$. Let C_R denote the point set common to all the point sets $S_R(z')$, where z' takes on all possible positions in D; the given set C surely belongs to C_R. It can be proved as in connection with Lemma II that C lies *interior* to C_R, and that C_R is closed.

LEMMA III. *If $r_n(z)$ is a rational function of z of degree n whose poles lie in D and if we have*

$$| r_n(z) | \leqq M, \qquad z \text{ on } C, \tag{1}$$

then we have also

$$| r_n(z) | \leqq MR^n, \qquad z \text{ on } C_R. \tag{2}$$

The proof of Lemma III is slightly more complicated than that of Lemma I because of the fact that in the present case the point set C may separate the plane. Let $z_1, z_2, \cdots, z_n$ be the poles of $r_n(z)$, not necessarily all distinct; if there are fewer than n poles, obvious modifications are to be made in our discussion. Let $z_1, z_2, \cdots, z_m$ be the poles of $r_n(z)$ which lie in $S(z_1)$. Denote by $\phi_1(z), \phi_2(z), \cdots, \phi_m(z)$ functions which map the region $S(z_1)$ onto the exterior of $\gamma: | w | = 1$ so that the respective points $z_1, z_2, \cdots, z_m$ correspond to $w = \infty$; these functions need not be single-valued, but that does not affect our reasoning. The function

* Compare for instance Walsh, *Crelle's Journal*, Vol. 159 (1928), pp. 197-209.

$$r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_m(z)$$

is analytic in $S(z_1)$ except possibly for branch points, and its modulus is single valued in $S(z_1)$. For z on C, and hence for z on the boundary of $S(z_1)$ inequality (1) is valid, so for z on the boundary of $S(z_1)$ we have

$$| r_n(z)/\phi_1(z)\phi_2(z) \cdots \phi_m(z) | \leqq M. \tag{10}$$

To be sure, $\phi_i(z)$ is not properly defined on the boundary of $S(z_1)$, but as z approaches that boundary, $| \phi_i(z) |$ approaches unity and $r_n(z)$ is continuous; inequality (10) is to be interpreted in that sense. It follows that (10) is valid throughout the region $S(z_1)$, a region within which the function on the left is analytic except possibly for branch points. Throughout the region common to $S(z_1)$ and C_R we have

$$| \phi_i(z) | \leqq R,$$

so from (10) we derive (2) for z in the region common to $S(z_1)$ and C_R. In a similar manner it is shown that (2) is valid for z in the region common to C_R and *any* $S(z_j)$ containing poles of $r_n(z)$.

Inequality (2) is valid on the boundary of each region $S(z_j)$ containing poles of $r_n(z)$; denote by $\sum$ the set complementary to the totality of such regions $S(z_j)$. Then inequality (2) is valid at each boundary point of $\sum$; the function $\phi_n(z)$ is analytic throughout $\sum$, so inequality (2) is valid at each point of $\sum$, and is valid in particular at each point of C_R belonging to $\sum$. We have already proved the validity of (2) at each point of C_R not in $\sum$, that is at each point of C_R lying in a region $S(z_j)$ containing poles of $r_n(z)$, so the proof of Lemma III is complete.

We add without proof the remarks * that if R' is less than R, the set $C_{R'}$ is interior to the set C_R; if E is an arbitrary closed set interior to C_R, there exists an $R' < R$ such that $C_{R'}$ contains E in its interior; if P represents an arbitrary point interior to C_R, there exists a value of R' less than R such that P lies on the boundary of $C_{R'}$.

Lemma III can now be applied to establish a result analogous to Theorem I; here C is no longer a Jordan region. In the present case *C and D are arbitrary closed point sets with no common point, such that the Dirichlet problem* (for arbitrary continuous boundary values) *has a solution for every region $S(z')$* defined as the totality of points which can be joined to a given point z' of D by broken lines not meeting C.

THEOREM II. *Let $r_n(z)$ be a sequence of rational functions of respective degrees n such that we have for z on C*

$$| f(z) - r_n(z) | \leqq M/\rho^n, \qquad \rho > 1. \tag{8}$$

* In the proof of these remarks it may be found convenient to use the results of Lebesgue, *Palermo Rendiconti*, Vol. 24 (1907), pp. 371-402.

If the poles of the functions $r_n(z)$ lie in D, then the sequence $r_n(z)$ converges uniformly for z on C_R, where $R < \rho^{1/2}$.

If (8) is valid for z in C and if the function $r_{n+1}(z) - r_n(z)$ is of degree $n+1$, with its poles in D, then the sequence $r_n(z)$ converges uniformly for z in C_R, where $R < \rho$.

If (8) is valid for z in C and if there exists a sequence $r'_n(z)$ of functions such that $r'_n(z) - r_n(z)$ is a rational function of degree n with its poles in D, if we have

$$|f(z) - r'_n(z)| \leqq M/\rho^n, \qquad z \text{ in } C, \tag{9}$$

and if the sequence $r'_n(z)$ converges uniformly for z in C_R, where $R < \rho$, [R may or may not depend on the sequence $r'_n(z)$], then the sequence $r_n(z)$ also converges uniformly for z in C_R.

The proof of Theorem II follows directly that of Theorem I.

III.

If the point set C of Lemma III actually separates the plane, and if the various regions into which C separates the plane are known to contain respectively specific numbers of poles of $r_n(z)$ less than n, then it may occur that a sharper result than Lemma III can be established, with a corresponding application more general than Theorem II. We give only a very simple illustration of this remark:

LEMMA IV. *Let C be an arbitrary Jordan curve in whose interior the origin lies. If $r_{2n}(z)$ is a rational function of z of the form*

$$r_{2n}(z) = a_{-n}z^{-n} + a_{-n+1}z^{-n+1} + \cdots + a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

and if we have for z on C

$$|r_{2n}(z)| \leqq M, \tag{11}$$

then we have also

$$|r_{2n}(z)| \leqq MR^n$$

for z on C_R, where D consists of the origin and the point at infinity.

The proof is essentially contained in the discussion already given in connection with Lemma III.

Lemma IV will be used in proving a result related to Theorem I:

THEOREM III. *Let C be an arbitrary Jordan curve in whose interior the origin lies. Let $r_{2n}(z)$ be a sequence of rational functions of the form*

$$\begin{aligned} r_{2n}(z) &= r'_n(z) + r_n''(z), \\ r'_n(z) &= a_{0n} + a_{1n}z + a_{2n}z^2 + \cdots + a_{nn}z^n, \\ r_n''(z) &= a_{-1n}z^{-1} + a_{-2n}z^{-2} + \cdots + a_{-nn}z^{-n}, \end{aligned}$$

such that we have for z on C

$$|f(z) - r_{2n}(z)| \leqq M/\rho^n, \qquad \rho > 1. \tag{12}$$

Then the sequence $r_{2n}(z)$ converges for z in C_R, uniformly for z in C_{R_1}, $R_1 < R$, where $R = \rho$. The point set D is to be taken as consisting of the origin and the point at infinity.

The region C_R is bounded by two curves, $C'_R = C_R(\infty)$ exterior to C and $C_R'' = C_R(0)$ interior to C; the sequence $r'_n(z)$ converges for z interior to C'_ρ, uniformly for z interior to C'_σ, $\sigma < \rho = R$, and the sequence $r_n''(z)$ converges for z exterior to C_ρ'', uniformly for z exterior to C_σ'', $\sigma < \rho = R$.

The first part of Theorem III follows directly, for we have by (12) for z *on* C

$$|f(z) - r_{2n}(z)| \leqq M/\rho^n, \qquad |f(z) - r_{2n+2}(z)| \leqq M/\rho^{n+1},$$
$$|r_{2n+2}(z) - r_{2n}(z)| \leqq (M/\rho^n)(1 + 1/\rho);$$

and this last inequality yields by Lemma IV, since the function

$$r_{2n+2}(z) - r_{2n}(z)$$

is of the form used in Lemma IV and of degree $2n + 2$,

$$|r_{2n+2}(z) - r_{2n}(z)| \leqq (M/\rho^n)(1 + 1/\rho) R_1^{n+1}$$

for z on C_{R_1}. Thus the sequence $r_{2n}(z)$ converges uniformly for z on C_{R_1} if $R_1 < \rho$.

The limit in C_ρ $(\rho = R)$ of the sequence $r_{2n}(z)$ is in C_ρ an analytic function of z, and coincides on C with $f(z)$, so this limit is the analytic extension in C_ρ of $f(z)$ and will be denoted by $f(z)$. The function $f(z)$, defined in the ring bounded by C'_ρ and C_ρ'', may be written in the form

$$f(z) = f_1(z) + f_2(z),$$

where $f_1(z)$ is analytic interior to C'_ρ and $f_2(z)$ is analytic exterior to C_ρ'' and vanishes at infinity: we have

$$f_1(z) = \frac{1}{2\pi i}\int_{C'_\sigma} \frac{f(t)\,dt}{t - z}, \quad \sigma < \rho, \qquad z \text{ interior to } C'_\sigma,$$
$$f_2(z) = \frac{1}{2\pi i}\int_{C_\sigma''} \frac{f(t)\,dt}{t - z}, \qquad z \text{ exterior to } C_\sigma''; \tag{13}$$

these integrals are independent of the precise value of σ. In a similar way we have

$$r'_n(z) = \frac{1}{2\pi i}\int_{C'_\sigma} \frac{r_{2n}(t)\,dt}{t - z}, \qquad z \text{ interior to } C'_\sigma,$$
$$r_n''(z) = \frac{1}{2\pi i}\int_{C_\sigma''} \frac{r_{2n}(t)\,dt}{t - z}, \qquad z \text{ exterior to } C_\sigma''. \tag{14}$$

We have already proved the uniform convergence of $r_{2n}(z)$ to $f(z)$ on C'_σ and C_σ''. Term-by-term integration of the equations just written yields

$$\lim_{n\to\infty} r'_n(z) = f_1(z), \quad \text{uniformly for } z \text{ interior to } C'_\tau, \quad \tau < \sigma < \rho,$$

$$\lim_{n\to\infty} r_n''(z) = f_2(z), \quad \text{uniformly for } z \text{ exterior to } C_\tau''.$$

Theorem III is now completely proved.

Theorem III yields a new result which the reader can readily formulate, by the use of a linear transformation $z = (z' - a)/(z' - b)$.

Theorem III, together with much more general results, has been proved by de la Vallée Poussin * for the case that C is a circle. Approximation on the unit circle by rational functions of the kind indicated is equivalent to approximation by trigonometric sums. Theorem III was previously proved by the present writer (*loc. cit.*) for the case that C is rectifiable; the previous proof is not valid in the present case. We recall also the following complement to Theorem III, which holds whether C is rectifiable or not:

If the function $f(z)$ is analytic in the closed region C_R, then there exist rational functions $r_{2n}(z)$ of the form indicated such that we have (12) *satisfied for $\rho = R$ for z on C.*

We have stated Theorem III for the case that C is a Jordan curve. The reader will notice that Lemma IV, Theorem III, and its complement are still true if C is an arbitrary closed set which separates the plane into two simply connected regions containing respectively the points $z = 0$ and $z = \infty$. Moreover, if we have $\rho > \sigma > 1$, there exists M' such that

$$| f(z) - r_{2n}(z) | \leqq M'\sigma^n/\rho^n, \qquad z \text{ on } C_\sigma.$$

Consequently we have by the use of equations (13) and (14)

$$| f_1(z) - r'_n(z) | \leqq M''\sigma^n/\rho^n, \qquad z \text{ interior to } C'_\tau \qquad \tau < \sigma < \rho,$$
$$| f_2(z) - r_n''(z) | \leqq M''\sigma^n/\rho^n, \qquad z \text{ exterior to } C_\tau''.$$

In particular we have

$$| f_1(z) - r'_n(z) | \leqq M''/\rho_1{}^n, \qquad z \text{ on } C,$$
$$| f_2(z) - r_n''(z) | \leqq M''/\rho_1{}^n, \qquad z \text{ on } C,$$

where ρ_1 is an arbitrary number less than ρ and where M'' depends on ρ_1.

IV.

There are various generalizations of our results which can be obtained with little difficulty. We merely mention these and leave the details to the reader.

---

* *Approximation des fonctions* (Paris, 1919), Ch. VIII.

(1) Functions $f(z)$ which are meromorphic instead of analytic on C may be approximated, where some poles of the approximating functions $r_n(z)$ are allowed on C, but with the requirement that the limit points of poles of the functions $r_{n+1}(z) - r_n(z)$ in Theorems I and II and of $r_{2n+2}(z) - r_{2n}(z)$ in Theorem III should lie in D. The results are entirely analogous to those already obtained.

(2) The properties of a sequence $r_n(z)$ and its limit often depend not primarily on the position of the poles of $r_n(z)$ but on the limiting position of the poles of $r_{n+1}(z) - r_n(z)$, that is on the limit points of those poles. A correspondingly more general statement of our results can be formulated.

(3) As has already been suggested, Lemma IV and Theorem III are merely special cases of more general results, where C separates the plane or not and D consists of several separated parts. Results exist more general than Lemma III and Theorem II, if the various parts of D are known to contain respectively a specific number (less than n) of poles of $r_n(z)$ or of $r_{n+1}(z) - r_n(z)$. This is true, moreover, whether or not parts of D are separated by C.

(4) Throughout the present paper our measure of the approximation of $r_n(z)$ to $f(z)$ has been

$$\max | f(z) - r_n(z) | , \qquad z \text{ on } C.$$

There exist other interesting measures of the approximation of $r_n(z)$ to $f(z)$ on C, such as (i) the line integral

$$\int | f(z) - r_n(z) |^p \ | dz | , \qquad p > 0,$$

extended over the boundary of C, if that boundary is rectifiable; (ii) the surface integral

$$\iint | f(z) - r_n(z) |^p \ dS, \qquad p > 0,$$

extended over the area of C; (iii) the integral

$$\int | f(z) - r_n(z) |^p \ | dw | , \qquad p > 0,$$

extended over the circle $\gamma: | w | = 1$ after mapping of C or of such a region as $S(z')$ onto the interior or exterior of γ. Any of these new measures of approximation may be used, even with the introduction of a positive weight or norm function, and under suitable restrictions there can be proved results analogous to those we have here established in detail.

(5) The approximation of harmonic functions by harmonic rational functions may be studied by our present methods.

ON APPROXIMATION TO A MAPPING FUNCTION BY POLYNOMIALS.

By O. J. FARRELL.*

1. *Introduction.* In a paper on approximation to an arbitrary function of a complex variable by polynomials, Professor Walsh † raised, among others, the question: *What conformal maps can be approximately performed by means of polynomials with an error uniformly and arbitrarily small?* The present paper contributes a new result in connection with this question.

We confine attention to the conformal mapping of a finite simply connected region on the interior of the unit circle. For such a map the question raised is answered by the following theorem:

THEOREM I. *Let G be a finite simply connected region in the z-plane, and let $f(z)$ be a function which maps G conformally on the interior of the unit circle. Then in order that $f(z)$ when suitably defined on the boundary γ of G be continuous on $G+\gamma$ and be also uniformly developable on $G+\gamma$ in a series of polynomials in z, it is necessary and sufficient that every point of γ be a simple ‡ boundary point of G and that γ be also the boundary of an infinite region.*

Theorem I is the chief result of the present paper; most of the paper is devoted to its proof. One further result, obtained by a simple application of Theorem I, may be mentioned here:

THEOREM II. *In the z-plane let G be a finite simply connected region whose boundary γ consists wholly of simple boundary points of G and is also the boundary, not of an infinite region, but of a finite region R distinct from G. Let $f(z)$ be a function which maps G conformally on the interior of the unit circle. Then $f(z)$ may be uniformly expanded on $G+\gamma$ in a series of polynomials in $1/(z-\alpha)$, where $z=\alpha$ is any chosen fixed point in R.*

An example of a region to which Theorem II is applicable is the region

* National Research Fellow.

† *Transactions of the American Mathematical Society*, Vol. 30 (1928), pp. 472-482.

‡ That is, contained in just one boundary element (Primende). See Carathéodory, "Über die Begrenzung einfach zusammenhängender Gebiete," *Mathematische Annalen*, Vol. 73 (1913), pp. 323-370, § 44.

consisting of a strip closed at one end which lies interior to the unit circle and which winds around infinitely often, approaching every point of the circle as it does so.

It is interesting to note in this connection the existence of a region G whose boundary γ although made up wholly of simple boundary points is not the boundary of any other region either finite or infinite. Consider, for instance, the region constructed by Professor Carathéodory * to show that boundary elements of the second kind can occur with the power of the continuum in every interval of boundary elements. Let P be any boundary point of this region that lies in the interior of the rectangle $ABCD$. Then it is readily seen that in a sufficiently small neighborhood of P are to be found only points of $G + \gamma$. This means of course that the complete boundary γ of G does not serve as the boundary of any region other than G.

2. *Proof of necessity of conditions of Theorem I.* The fact that the boundary γ of G must consist entirely of simple boundary points of G results from the following

LEMMA. *Let $w = f(z)$ be a function which maps a finite simply connected region G of the z-plane conformally on the region $|w| < 1$. A necessary and sufficient condition that $f(z)$ be definable on the boundary γ of G so as to be continuous on $G + \gamma$ is that every point of γ be a simple boundary point of G.*

The proof of this lemma is immediate by Theorems XIII and XXI of the paper by Carathéodory already referred to above.

It remains to show that γ must also be the boundary of an infinite region. Suppose this were not so. All points of the plane which can be joined with the point $z = \infty$ by a Jordan arc containing no point of γ form a region whose boundary β consists wholly of points of γ, but which by hypothesis does not exhaust the point set γ. Let z_0 be any fixed point of G. All points which can be joined to z_0 by a Jordan arc containing no points of β form a region T bounded by β and containing as interior points those points of γ which do not belong to β.

If now there exists a sequence of polynomials converging uniformly to $f(z)$ on $G + \gamma$, the sequence will in particular converge uniformly on β and hence also on $T + \beta$. This sequence thus defines a function regular-analytic in T, continuous on $T + \beta$, and assuming its maximum absolute value, namely unity, at interior points of T. But this is impossible. Hence

* *Loc. cit.*, § 49.

the complete boundary γ of G must also constitute the boundary of an infinite region.

3. *A theorem on the convergence of a sequence of mapping functions.* We require for the proof of the sufficiency of the conditions of Theorem I the following theorem:

THEOREM III. *In the z-plane let G be a finite simply connected region containing the origin whose boundary γ consists wholly of simple boundary points of G and is also the boundary of an infinite region. Let*

$$G_1, G_2, G_3, \cdots G_n, \cdots \tag{1}$$

denote a sequence of uniformly bounded simply connected regions such that the closed region $G+\gamma$ lies interior to every G_n. Let $w=f_1(z)$, $f_2(z)$, $f_3(z), \cdots$; $f(z)$ denote the functions that map conformally $G_1, G_2, G_3, \cdots$; G respectively on the region $|w|<1$ so that in every case the points $z=0$ and $w=0$ together with the directions of the positive axes of reals at these points correspond to each other. Suppose that $f(z)$ has been defined on γ so as to be continuous on $G+\gamma$. Then if G be the kernel of the sequence (1) and if this sequence converge to its kernel, we have*

$$\lim_{n\to\infty} f_n(z) = f(z) \tag{2}$$

uniformly for all z on $G+\gamma$.

It follows from the hypotheses of Theorem III by results of Bieberbach† that (2) holds uniformly in any closed region lying wholly interior to G. We shall show, however, that (2) holds uniformly for all z in the open region G. From this it will then follow, due to the continuity of $f(z)$ and $f_n(z)$ on $G+\gamma$, that (2) holds uniformly for all z on $G+\gamma$.

Suppose that (2) does not hold uniformly for all z in the open region G. Then there exists a sequence of points of G

$$z_1, z_2, \cdots z_k, \cdots \tag{3}$$

together with a subsequence of functions

$$f_{\mu_1}(z), f_{\mu_2}(z), \cdots f_{\mu_k}(z), \cdots \tag{4}$$

and a positive number Δ such that

$$|f_{\mu_k}(z_k) - f(z_k)| > \Delta, \qquad (k=1,2,3,\cdots). \tag{5}$$

---

* See Bieberbach, *Lehrbuch der Funktionentheorie*, Vol. 2 (1927), pp. 12-13.

† *Loc. cit.*, p. 13. Compare Walsh, *loc. cit.*, p. 474.

No interior point of G can be a limit point of (3) because of the uniform convergence of $f_n(z)$ in every closed region that lies in G. We may assume, then, that (3) converges in the sense of Carathéodory * to a boundary element E_0 of G, since in any event a subsequence can be chosen which has this property. Let w_0 designate the point on the circle $|w| = 1$ which corresponds by $w = f(z)$ to E_0.

The boundary element E_0 may be defined by a chain of cross-cuts $q_1, q_2, \cdots$ which lie on concentric circles and converge to a point of E_0. Let $g_1, g_2, \cdots$ denote the corresponding chain of subregions.

A point in G_n which can be connected to a point on the boundary γ_n of G_n by a Jordan arc J of diameter less than δ is transformed by $w = f_n(z)$ into a point whose distance from the point corresponding by this same relation to the boundary element of G_n defined by J is less than a quantity $\theta(\delta)$ which is independent of n and approaches zero with δ.†

Let an arbitrary positive ϵ be given. Choose δ so that $\theta(\delta) < \epsilon$. From the cross-cuts $q_1, q_2, \cdots$ choose one q_N whose length is less that $\delta/3$, whose transform by $w = f(z)$ lies entirely within distance ϵ of w_0, and such that the point $z = 0$ is exterior to g_N. Let z' designate the middle point of q_N. Denote by Q_1 and Q_2 the end-points ‡ of q_N. Choose a positive number η less than $\delta/3$ and also less than $\overline{Q_1Q_2}/2$.

By choosing n_0 sufficiently large, we have for all $n \geqq n_0$

(a) $|f_n(z') - f(z')| < \epsilon$,

(b) within distance η of every point of γ lie points of γ_n.

The condition (a) can be obtained by virtue of the result of Bieberbach mentioned above. The fact that (b) can also be obtained may be seen as follows. If this were not possible, there would exist a sequence of points $P_1, P_2, P_3, \cdots$ on γ together with a subsequence of regions $G_{\mu_1}, G_{\mu_2}, G_{\mu_3}, \cdots$ such that for every k no point of the boundary γ_{μ_k} of G_{μ_k} lies within distance η of P_k. We may assume that the sequence $P_1, P_2, \cdots$ has a unique limit point P_0. Then all points P_k for sufficiently large k lie within distance $\eta/2$ of P_0. Hence no boundary γ_{μ_k} from a certain k on comes within distance $\eta/2$ of P_0.

---

* *Loc. cit.*, definition No. III, top of p. 332; also Theorem VII, p. 341.

† See Lindelöf, "Sur un principe général de l'analyse," *Acta Societatis Scientiarum Fennicae*, Vol. 46, No. 4, the two italicized results on page 16. Compare Walsh, *loc. cit.*, p. 475.

‡ The points Q_1 and Q_2 are distinct. If this were not so, points of γ and hence also points of the infinite region F bounded by γ could be found both inside and outside the circle formed by q_N. But this would be absurd, for F would not then be a region.

But then G would not be the kernel of the subsequence $G_{\mu_1}, G_{\mu_2}, \cdots$, contrary to the hypothesis that the sequence (1) converges to its kernel G.

Consider any region G_n of (1) for which $n \geqq n_0$. Join Q_1 and Q_2 each with the nearest point of γ_n by means of the straight line segments $\overline{Q_1T_1}$ and $\overline{Q_2T_2}$. Let R_1 and R_2 denote the first points of intersection of $\overline{Q_1T_1}$ and $\overline{Q_2T_2}$ respectively with q_N when one proceeds from T_1 to Q_1 and from T_2 to Q_2. (It can happen of course in simple cases that R_1 coincides with Q_1, or R_2 with Q_2, or both.) Thus we obtain a cross-cut $T_1R_1R_2T_2$ of G_n consisting of the straight line segments $\overline{T_1R_1}$ and $\overline{R_2T_2}$ together with that portion of q_N which joins R_1 and R_2. Call this cross-cut σ. We notice that σ contains z' and that the length of σ is less than δ.

Denote by G'_n and G_n'' the two sub-regions of G_n defined by σ. For sufficiently small choice of η the subregion g_{N+1} of G is wholly interior to one of the two regions G'_n and G_n'', say G'_n, while the point $z = 0$ is interior to the other region G_n''. This may be seen as follows. About z' as center draw a circle K lying entirely interior to G. The cross-cut q_N divides the interior of K into two regions H' and H'' such that one of them, say H', is entirely interior to g_N and H'' is wholly exterior to g_N. Join an arbitrary fixed point of H' with an arbitrary fixed point of g_{N+1} by a Jordan arc lying in g_N. Join also a fixed point of H'' with the point $z = 0$ by a Jordan arc lying in G but exterior to g_N. Let μ denote the minimum distance between a point of either arc and the boundary γ of G. Let ρ denote the minimum distance between either of the points Q_1, Q_2 and a point of K. Finally let λ denote the minimum distance * between q_N and the closed region consisting of g_{N+1} and its boundary. Then it will be seen that we need merely to choose any positive η which (in addition to being less than $\delta/3$ and $\overline{Q_1Q_2}/2$ as already required above) is less than the least of the positive numbers μ, ρ, λ.

The transform τ of σ by $w = f_n(z)$ is a cross-cut of the region $|w| < 1$ lying entirely within distance $\theta(\delta) < \epsilon$ of the point on the circle $|w| = 1$ which corresponds by $w = f_n(z)$ to the boundary element of G_n defined by the segment $\overline{Q_1T_1}$. Thus every point of τ is within distance 2ϵ of the point $w = f_n(z')$. It follows then from the inequalities

$$|f_n(z') - f(z')| < \epsilon \quad \text{and} \quad |f(z') - w_0| < \epsilon$$

that all points of τ are within distance 4ϵ of w_0. Hence one of the two sub-regions of $|w| < 1$ defined by τ must lie wholly within distance 4ϵ of w_0.

* This minimum distance λ is positive. In the contrary case, an end-point of q_N, say Q_1, would have to lie on the boundary of g_{N+1}. But then Q_1 would not be a simple boundary point of G. See Carathéodory, *loc. cit.*, Theorem XXII.

And this region is the transform by $w = f_n(z)$ of G'_n, at least for $\epsilon < \frac{1}{4}$, since then the point $w = 0$, which corresponds to $z = 0$, is a point of the transform of the other region G_n''.

The transform of g_{N+1} by $w = f_n(z)$ is therefore a region every point of which is within distance 4ϵ of w_0. And the transform of g_{N+1} by $w = f(z)$ lies wholly within distance ϵ of w_0, since g_{N+1} is a part of g_N. And so we have for $\epsilon < 5\Delta$

$$| f_n(z) - f(z) | < \Delta$$

for all z in g_{N+1} and for all $n \geqq n_0$. But for sufficiently large k the points z_k of (3) are interior points of g_{N+1}. And for these points (5) holds. Here we have a contradiction, and hence it must be true that (2) holds uniformly in the open region G and consequently also on $G + \gamma$.

4. *Proof of sufficiency of conditions of Theorem I.* It is no loss of generality to assume that the given region G contains the origin and that the mapping makes correspond the points $z = 0$ and $w = f(z) = 0$ together with the directions of the positive axes of reals at these points; for this may always be effected by a preliminary linear transformation of the form

$$\zeta = az + b. \tag{6}$$

Then $f(z)$ becomes $g(\zeta)$, and an expansion of $g(\zeta)$ in the transform of $G + \gamma$ by (6) in terms of polynomials in ζ is equivalent to an expansion of $f(z)$ on $G + \gamma$ in terms of polynomials in $az + b$, that is, in terms of polynomials in z.

Choose a sequence of uniformly bounded simply connected regions $G_1, G_2, \cdots$ which have the given region G as their kernel and which converge to their kernel. Let $w = f_1(z), f_2(z), \cdots$ denote the functions which map these regions respectively on the interior of the unit circle so that in every instance the points $z = 0$ and $w = 0$ correspond, together with the directions of the positive axes of reals at these points.

Let an arbitrary positive ϵ be given. Choose by virtue of Theorem III a function $f_n(z)$ such that

$$| f(z) - f_n(z) | < \epsilon/2, \qquad z \text{ on } G + \gamma. \tag{7}$$

By Runge's well known theorem * there exists a polynomial $P(z)$ such that

$$| f_n(z) - P(z) | < \epsilon/2, \qquad z \text{ on } G + \gamma. \tag{8}$$

Combining (7) and (8) we have the desired result

* See, for instance, Bieberbach, *loc. cit.*, Vol. 1 (1921), p. 296.

$$| f(z) - P(z) | < \epsilon, \qquad z \text{ on } G + \gamma.$$

5. *Two applications of Theorem I.* Our first application of Theorem I is Theorem II, the proof of which is very simple. Let $z = \alpha$ be an arbitrarily chosen fixed point in R, and let the circle C: $| z - \alpha | = r$ lie wholly in R. The transformation

$$\zeta = r^2/(z - \alpha) + \alpha \tag{9}$$

exchanges the interior of C with its exterior. Let $g(\zeta)$ denote the inverse of (9). Then by Theorem III the function $f[g(\zeta)]$, which maps the transform of G by (9) on the region $| w | < 1$, may be uniformly approximated as closely as desired in this closed region by a polynomial in ζ. But this means that $f(z)$ can be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in $1/(z - \alpha)$.

Another simple application of Theorem I results in

THEOREM IV. *In the z-plane let G be a region whose boundary γ consists wholly of simple boundary points of G and is also the boundary of an infinite region. If the function $F(z)$ is analytic in G, continuous on $G + \gamma$, and constant on each boundary element of G, then $F(z)$ may be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in z.*

For convenience in proving this theorem we require the following

LEMMA. *The function $f^n(z)$, where n is any positive integer and $f(z)$ is the mapping function of Theorem I, may be uniformly approximated as closely as desired on $G + \gamma$ by a polynomial in z.*

Let an arbitrary positive ϵ be assigned. Choose by virtue of Theorem I a polynomial $P(z)$ such that

$$| P(z) - f(z) | < \frac{\epsilon}{n(1+\epsilon)^{n-1}}, \qquad z \text{ on } G + \gamma.$$

Then we have

$$\begin{aligned} | P^n(z) - f^n(z) | &= | P(z) - f(z) | \\ &\quad \cdot | P^{n-1}(z) + P^{n-2}(z)f(z) + P^{n-3}(z)f^2(z) + \cdots + f^{n-1}(z) | \\ &< \frac{\epsilon}{n(1+\epsilon)^{n-1}} \cdot n(1+\epsilon)^{n-1} = \epsilon. \end{aligned}$$

Let $z = g(w)$ be the inverse of $w = f(z)$, where $f(z)$ is a function which maps the given region G of Theorem IV on the region $| w | < 1$. Then for $| w | < 1$ we have

$$F(z) = F[g(w)] = H(w),$$

and $H(w)$ is analytic for $|w| < 1$. Moreover, $H(w)$ takes on a boundary value at each point of the circle $|w| = 1$. Let $w = w_0$, for instance, be a point of this circle, and let E_0 denote the boundary element of G that corresponds by $z = g(w)$ to w_0. If $w_1, w_2, \cdots$ be any sequence of points interior to $|w| < 1$ converging to w_0, the corresponding sequence $z_1, z_2, \cdots$ $(z_n = g(w_n))$ converges to E_0. But then the sequence

$$H(w_1), H(w_2), \cdots,$$

or what is the same,

$$F(z_1), F(z_2), \cdots,$$

converges to a definite value, namely the constant value of $F(z)$ on E_0. So if we define $H(w)$ on $|w| = 1$ as equal to its boundary values there, $H(w)$ is continuous for $|w| \leqq 1$ as well as analytic for $|w| < 1$.

Then, given an arbitrary positive ϵ, we can find a polynomial * $P(w)$ such that

$$|H(w) - P(w)| < \epsilon/2, \qquad |w| \leqq 1.$$

That is

$$|F(z) - P[f(z)]| < \epsilon/2, \qquad z \text{ on } G + \gamma.$$

But by the lemma each positive integral power of $f(z)$ can be approximated as closely as desired on $G + \gamma$ by a polynomial in z. Consequently a polynomial $p(z)$ exists such that

$$|P[f(z)] - p(z)| < \epsilon/2, \qquad z \text{ on } G + \gamma.$$

Combining the last two inequalities we have

$$|F(z) - p(z)| < \epsilon, \qquad z \text{ on } G + \gamma.$$

UNIVERSITY OF MUNICH,
MUNICH, GERMANY.

* See Walsh, *loc. cit.*, Theorem I.

ON INFINITE ORTHOGONAL MATRICES.

By MONROE HARNISH MARTIN.

TABLE OF CONTENTS.

INTRODUCTION.

I. Reality discussion of the trigonometrical moment problem of Herglotz.

II. Application of the trigonometrical moment problem to orthogonal matrices.

III. A connection between reduced skew-symmetric and reduced Hermitian matrices.

IV. The parametric representation of the manifold of the orthogonal matrices.

V. The four classes of orthogonal matrices.

VI. Rotations and reflections.

APPENDIX. The problem of moments for a finite interval and the theory of bounded Hermitian matrices.

INTRODUCTION.

If one interprets the elements m_{jk} $(j, k = 1, 2, \cdots n < +\infty)$ of an arbitrary real or complex n rowed square matrix $\boldsymbol{M} = \| m_{jk} \|$ as the coördinates of a point in a Euclidean space of n^2 real or complex dimensions, the set of points representing the real unitary matrices (i. e. the orthogonal matrices) divides itself into two separate continua, one of which contains the rotation matrices (i. e. the orthogonal matrices of determinant $+1$) and the other the reflection matrices (i. e. the orthogonal matrices of determinant -1). The interpretation of a finite matrix as a point in a Euclidean space of a suitable number of dimensions may obviously be extended to infinite matrices by employing an abstract manifold of ∞^2 real or complex dimensions. The aim of the present paper is to obtain a division of this abstract manifold into two sub-manifolds in a manner which is a natural extension of the afore-mentioned division of the space of finite orthogonal matrices. The procedure is to first obtain a one-to-one parametric representation of the entire manifold of the infinite orthogonal matrices which divides the ortho-

gonal matrices into four distinct classes, and then to show how the matrices of these four classes may be reapportioned into two sub-manifolds of the manifold of the infinite orthogonal matrices, the two sub-manifolds coinciding, for finite orthogonal matrices, with the two separate continua noted above. Since the formula of Cayley does not yield, as we show in some detail in the next few pages, a parametric representation of the entire manifold of the orthogonal matrices, it has been found necessary to employ the spectral theory which yields a one-to-one exponential representation of the entire manifold of the orthogonal matrices. Throughout this paper we employ bold-faced capital letters to designate finite matrices, the Roman capitals being retained for infinite matrices.

We denote by

$$\Phi[\mathrm{M};x]=\sum_{p=1}^{+\infty}\sum_{q=1}^{+\infty}m_{pq}x_p\bar{x}_q, \tag{1}$$

the *compound form* (Kopplungsform) of an arbitrary bounded † matrix $\mathrm{M}=\|\,m_{pq}\,\|$ and shall call a real or complex vector $x=(x_1,x_2,\cdots)$ a *unit vector* if

$$|\,x\,|=(\Phi[\mathrm{E};x])^{\frac{1}{2}}=(\sum_{p=1}^{+\infty}|\,x_p\,|^2)^{\frac{1}{2}}=1, \tag{2}$$

in which E is the unit matrix. The compound forms of bounded matrices possess the obvious uniqueness property in that the necessary and sufficient condition for the equality

$$\Phi[\mathrm{M};x]=\Phi[\mathrm{N};x]$$

to hold for every unit vector x is that $\mathrm{M}=\mathrm{N}$. Expressed otherwise, the equality

$$\Phi[\mathrm{M};x]=0,$$

holds for every unit vector x then and only then if M is the zero matrix $\|\,0\,\|$.

We shall understand by the *domain* ‡ $\Delta(\mathrm{M})$ of an arbitrary bounded matrix M the point-set which one constructs in the complex plane from the values, together with their limiting values, which are taken by the compound form (1) for those vectors x satisfying (2), i. e. for the totality of the unit vectors which compose the so-called complex Hilbert sphere (2). The domain $\Delta(\mathrm{M})$ of a bounded matrix M is then, from its definition, a bounded closed point set in the complex plane.

† D. Hilbert, "Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen," (1912), p. IV. This book will be referred to as "Hi".

‡ Termed by German writers "Wertevorrat."

According to a theorem of Toeplitz † there exists for every bounded matrix M at most one bounded matrix K for which the two independent conditions $KM = E$, $MK = E$ are fulfilled. If such a bounded matrix exists we designate it as the reciprocal of M and denote it by M^{-1}. We then understand by the *spectrum* ‡ of a bounded matrix N the set of those points λ in the complex plane for which $(\lambda E - N)^{-1}$ does not exist and denote this point-set by $\Lambda(N)$. It is known § that the point-set $\Lambda(N)$ thus defined for every bounded matrix N contains at least one point and is a closed subset of the domain $\Delta(N)$. If we denote by $l[M]$ the greatest lower bound of the compound form $\Phi[N; x]$ of a bounded matrix N under the condition that $|x| = 1$ and by N^* the transposed matrix of the conjugate complex elements of N (i. e. $N^* = \bar{N}'$), the number λ is then and only then, according to Toeplitz,¶ in the spectrum of N if at least one of the following two equations is fulfilled

$$(3) \quad l[(\lambda E - N)(\lambda E - N)^*] = 0, \qquad l[(\lambda E - N)^*(\lambda E - N)] = 0.$$

If N is commutable with N^*, i. e. N is a normal matrix, the two conditions in (3) are identical.

The subset of the spectrum composed of those and only those points λ for which the homogeneous equations belonging to the bounded matrix $\lambda E - N$ possess a solution x for which $|x| = 1$ is called the *point spectrum* ‖ of N. The points of the spectrum do not necessarily all belong to the point spectrum and it is possible that the point spectrum of N contains no single point. It is clear that the spectrum of $cE - N$ is obtained from the spectrum of N by a translation, determined by the real or complex number c, in the complex plane. If T is a bounded matrix for which T^{-1} exists then the bounded matrices M and TMT^{-1} possess the same spectrum and the same point spectrum.

A matrix N is said to be *Hermitian* if $N^* = N$, *unitary* if N^{-1} exists and is equal to N^*, and *skew-symmetric* if $N' = -N$. It is known that the unitary matrices form a group of bounded matrices and that no unitary

† O. Toeplitz, "Die Jacobische Transformation der quadratischen Formen von unendlichvielen Veränderlichen," *Göttingen Nachrichten* (1907), pp. 101-109. This paper will be referred to as "T".

‡ A. Wintner, "Zur Theorie der beschränkten Bilinearformen," *Mathematische Zeitschrift*, Vol. 30 (1929), p. 239. This paper will be referred to as "W".

§ See "W", pp. 241-243.

¶ See "T".

‖ This definition of the point spectrum of an arbitrary bounded matrix N coincides for bounded Hermitian and unitary matrices with the usual definition for the point spectrum of matrices of these two classes.

matrix is completely continuous (vollstetig). The domain of an arbitrary bounded matrix N is a unitary invariant, i. e. if U is any unitary matrix, then $\Delta(N) = \Delta(UNU^{-1})$.

The domain $\Delta(H)$ of a bounded Hermitian matrix H is a bounded, closed interval of the real axis of the complex plane, the end-points of which always belong † to the spectrum $\Lambda(H)$ of H. Since the spectrum of any bounded matrix, as has been mentioned above, is a subset of the domain of the matrix, the spectrum $\Lambda(H)$ of a Hermitian matrix H lies entirely on the real axis of the complex plane. The domain of a unitary matrix U is a convex subset of the unit circle, the spectrum of U being contained on the boundary of the unit circle. The real subset of the manifold of the unitary matrices forms the manifold of the *orthogonal matrices* O so that the spectrum of O is likewise contained on the boundary of the unit circle and is in addition symmetric with respect to the axis of reals inasmuch as the spectrum of any real bounded matrix possesses this symmetry property. The domain of a bounded real skew-symmetric matrix is contained on the axis of imaginaries and, as may easily be verified directly, is symmetric with respect to the axis of reals. For our purpose it will be more instructive to derive this symmetry property of the domain of a bounded real skew-symmetric matrix as follows: The matrix iS in which S is a bounded real skew-symmetric matrix is obviously Hermitian so that its domain and, therefore, its spectrum is contained on the axis of reals. Since the spectrum of S is symmetric with respect to the axis of reals, the spectrum of iS is symmetric with respect to the axis of imaginaries and, therefore, also the domain of iS is symmetric with respect to the axis of imaginaries inasmuch as, as has been mentioned above, the end-points of the domain of a bounded Hermitian matrix H always belong to the spectrum of H. Finally, since $\Phi[S; x] = -i\Phi[iS; x]$, it follows that the domain of S is symmetric with respect to the axis of reals.

A pair of matrices (M, N) is said to be *separated* if the elements m_{pq} of M and the elements n_{pq} of N stand in the following relationship to one another:

$$(4) \qquad \begin{array}{lll} \text{if} & m_{pq} \neq 0 & \text{then} \quad n_{pk} = n_{jp} = n_{jq} = n_{qk} = 0, \\ \text{if} & n_{pq} \neq 0 & \text{then} \quad m_{pk} = m_{jp} = m_{jq} = m_{qk} = 0, \end{array}$$

in which the two subscripts j, k take independently (p and q fixed) all positive integral values.

---

† F. Riesz, "Les systèmes d'équations linéaires a une infinité d'inconnues" (1913), pp. 139-140. For a direct demonstration cf. A. Wintner, "Spektraltheorie der unendlichen Matrizen" (1929), pp. 146-148. The book of Riesz will be referred to as "R" and that of Wintner as "S".

If two bounded matrices M and N form a separated pair (M, N) and if Φ_M Φ_N are any two points whatsoever in the domains $\Delta(M)$ and $\Delta(N)$ respectively, it is clear from (4) that there exist two sequences of unit vectors so that (cf. the definition of the domain of a bounded matrix on p. 580)

$$(5)\quad \begin{aligned} \Phi_M &= \lim_{k\to+\infty} \Phi\,[M; y_k] \quad \text{while} \quad 0 = \Phi\,[N; y_k] \quad \text{for} \quad k = 1, 2, \cdots, \\ \Phi_N &= \lim_{k\to+\infty} \Phi\,[N; z_k] \quad \text{while} \quad 0 = \Phi\,[M; z_k] \quad \text{for} \quad k = 1, 2, \cdots, \end{aligned}$$

It is not difficult to see that if the matrix pair (M, N) is separated, the infinitely many matrix pairs

$$(6)\qquad (M^p, N^q); \qquad (p, q = 1, 2, \cdots)$$

are also separated. A pair of bounded matrices (M, N) is said to be *perpendicular* if

$$(7)\qquad MN = NM = \| 0 \|,$$

and it is easy to see that while a separated pair of bounded matrices is always perpendicular, it does not follow that a perpendicular pair of matrices is necessarily separated. In addition the property of a pair of bounded matrices to be perpendicular is readily seen, in contrast to the property of a matrix pair to be separated, to be invariant with respect to simultaneous matrix transformations of the two matrices composing the pair. It is trivial from (7) that for a perpendicular matrix pair (M, N) we have

$$(8)\qquad (M + N)^n = M^n + N^n; \qquad (n = 1, 2, \cdots).$$

We employ (5), (6) and (8) in order to show that a separated pair of Hermitian matrices (M, N) possesses in the notation of p. 581 the property

$$(9)\qquad \Lambda(M + N) = \Lambda(M) + \Lambda(N)$$

in which the + sign denotes the usual process of the addition of two point sets, i. e. for example $\Lambda(M) + \Lambda(M) = \Lambda(M)$. In order to prove (9) let μ and ν be any two points (not necessarily different) in the respective spectra of the bounded Hermitian matrices M and N forming the separated matrix pair (M, N). According to (3) there accordingly exist two sequences $\{y_n\}$ and $\{z_n\}$ of unit vectors for which

$$(10)\quad \begin{aligned} &\lim_{n\to+\infty} \Phi\,[(\mu E - M)(\mu E - M)^*; y_n] \\ &\qquad = \lim_{n\to+\infty} \Phi\,[(\mu^2 E - 2\mu M + M^2); y_n] = 0, \\ &\lim_{n\to+\infty} \Phi\,[(\nu E - N)(\nu E - N)^*; z_n] \\ &\qquad = \lim_{n\to+\infty} \Phi\,[(\nu^2 E - 2\nu N + N^2); z_n] = 0, \end{aligned}$$

and we now employ (5), (6) and (8) to show that both μ and ν lie in the spectrum of the Hermitian matrix $\mathrm{M}+\mathrm{N}$. In order that a point λ lies in the spectrum of $\mathrm{M}+\mathrm{N}$ it is, again according to (3), necessary and sufficient that there exist a sequence of unit vectors $\{x_n\}$ for which

$$(11)\quad \lim_{n\to+\infty} \Phi\,[(\lambda\mathrm{E}-\mathrm{M}-\mathrm{N})(\lambda\mathrm{E}-\mathrm{M}-\mathrm{N})^*;x_n] = \lim_{n\to+\infty} \Phi\,[(\lambda^2\mathrm{E}-2\lambda\mathrm{M}-2\lambda\mathrm{N}+\mathrm{M}^2+\mathrm{N}^2);x_n]=0,$$

inasmuch as from (8) we have $(\mathrm{M}+\mathrm{N})^2=\mathrm{M}^2+\mathrm{N}^2$. Since, according to (6), the matrix pairs (M,N), $(\mathrm{M},\mathrm{N}^2)$, $(\mathrm{M}^2,\mathrm{N})$, $(\mathrm{M}^2,\mathrm{N}^2)$ are all separated, the condition (11) may obviously, from (10), be satisfied by putting either

$$\lambda=\mu \text{ and } \{x_n\}=\{y_n\} \quad \text{or} \quad \lambda=\nu \text{ and } \{x_n\}=\{z_n\},$$

inasmuch as, as follows from (5) and (6), the sequences $\{y_n\}$ and $\{z_n\}$ in (10) may be chosen to have

$$\Phi\,[\mathrm{M};z_n]=0,\ \Phi\,[\mathrm{M}^2;z_n]=0;\quad \Phi\,[\mathrm{N};y_n]=0,\ \Phi\,[\mathrm{N}^2;y_n]=0,$$

for $n=1,2,\cdots$, i. e. the points μ and ν both belong to the spectrum of $\mathrm{M}+\mathrm{N}$.

An important application for p. 607 of the text concerns itself with the special case in which M is a diagonal matrix, the elements of which are either 0 or $+1$. The complementary matrix $\mathrm{E}-\mathrm{M}$ is accordingly of the same type as M and forms with M the separated pair $(\mathrm{E}-\mathrm{M},\mathrm{M})$. We then construct the separated pair $(c\,[\mathrm{E}-\mathrm{M}]+\mathrm{N},c\mathrm{M})$ where c is any number, real or complex, and apply the result (9) of the preceding paragraph to this matrix pair, thereby obtaining

$$\Lambda(c\mathrm{E}+\mathrm{N})=\Lambda(c\,[\mathrm{E}-\mathrm{M}]+\mathrm{N}+c\mathrm{M})=\Lambda(c[\mathrm{E}-\mathrm{M}]+\mathrm{N})+\Lambda(c\mathrm{M}),$$

or, in particular for $c=\pi$ and $\mathrm{N}=i\mathrm{S}$

$$(12)\qquad \Lambda(\pi\mathrm{M})+\Lambda(\pi\,[\mathrm{E}-\mathrm{M}]+i\mathrm{S})=\Lambda(\pi\mathrm{E}+i\mathrm{S}).$$

Since the spectrum $\Lambda(\pi\mathrm{M})$ of $\pi\mathrm{M}$ only contains the two points 0 and π, it is clear from (12) that $\lambda>\pi$ is then and only then a point in the spectrum of $\pi\,[\mathrm{E}-\mathrm{M}]+i\mathrm{S}$ if it is a point in the spectrum of $\pi\mathrm{E}+i\mathrm{S}$. It is furthermore trivial that 2π is then and only then in the point spectrum of $\pi\mathrm{E}+i\mathrm{S}$ if $+\pi$ (and if S is real and skew-symmetric then also $-\pi$) is a point in the point spectrum of $i\mathrm{S}$ so that we are led to the following result employed on p. 610 of the text: If S is a real skew-symmetric matrix whose spectrum is contained in the closed interval $[-i\pi,+i\pi]$ and if M is a diagonal matrix, the elements of which are either 0 or $+1$, forming with S the

separated matrix pair (M, S), then the point 2π lies in the point spectrum of the Hermitian matrix $\pi(E-M)+iS$ if, and only if, if the end-points $-\pi$ and $+\pi$ of the interval $[-\pi, +\pi]$ belong to the point spectrum of iS.

A matrix I is said to be an *idempotent* matrix if $I^2 = I$ (inasmuch as then $I^n = I$ for $n = 1, 2, \cdots$). The Hermitian idempotent matrices are designated by Hilbert as "Einzelmatrizen." The bounded idempotent matrices I possess the property that the continuously many matrices $2\pi in I$ ($n = 0, \pm 1, \cdots$) are all "logarithms" of E, i. e. $\exp(2\pi in I) = E$ [Cf. (80)]. Since the matrix

$$[(\lambda-1)E+I]/[\lambda(\lambda-1)]$$

represents, as may easily be verified, for $\lambda(\lambda-1) \neq 0$ the reciprocal matrix of $(\lambda E - I)$ for every bounded idempotent matrix I, the spectrum of I contains at most the two points $\lambda = 0$ and $\lambda = +1$. For real symmetric (and therefore bounded) idempotent matrices there exists, according to Hilbert,† an orthogonal matrix O so that $OIO^{-1} = \| \kappa_{pq} \|$ where $\kappa_{pq} = 0$ for $p \neq q$ and $\kappa_{pp} = 0$ or $\kappa_{pp} = +1$ (Cf. the matrices M and $E-M$ above). We denote by α' the number of times 0 occurs in the sequence $\{\kappa_{pp}\}$ and by β' the number of times $+1$ occurs in this sequence so that $\alpha' \geqq 0$, $\beta' \geqq 0$ and $\alpha' + \beta' = +\infty$. The numbers α' and β' are uniquely determined by the matrix I, i. e. they are independent of the special choice of the matrix O employed in transforming I into the diagonal form $\| \kappa_{pq} \|$. It will be necessary to distinguish the following four classes of idempotent matrices:

I:	$\alpha' = +\infty$,	$\beta' = 2m$	for $0 \leqq m < +\infty$,
II:	$\alpha' = +\infty$,	$\beta' = 2m+1$	for $0 \leqq m < +\infty$,
III:	$\alpha' = m$,	$\beta' = +\infty$	for $0 \leqq m < +\infty$,
IV:	$\alpha' = +\infty$,	$\beta' = +\infty$.	

We call a matrix Q a *square root of E* if $Q^2 = E$. The bounded square roots of E may be characterized by the matrix equation $Q^{-1} = Q$ and possess the property that the continuously many matrices $2\pi in Q$ ($n = 0, \pm 1, \pm 2, \cdots$) are all "logarithms" of E (Cf. above). The general real symmetric square root of E depends (as does the general real symmetric idempotent matrix) upon infinitely many arbitrary constants. Since $(E\lambda + Q)(\lambda^2 - 1)^{-1}$ represents, as may easily be verified, for $\lambda^2 - 1 \neq 0$ the reciprocal matrix of $\lambda E - Q$ for every bounded square root Q of E, the spectrum of Q contains at most the two points $\lambda = +1$ and $\lambda = -1$. It therefore follows from

† See "Hi", p. 145.

the theorem of Hilbert † on bounded real symmetric matrices without continuous spectra that there exists for every real symmetric (and therefore bounded) square root of E an orthogonal matrix O so that $OQO^{-1} = \| \epsilon_{pq} \|$ where $\epsilon_{pq} = 0$ for $p \neq q$ and $\epsilon_{pp} = +1$ or $\epsilon_{pp} = -1$. We denote by α the number of times $+1$ occurs in the sequence $\{\epsilon_{pp}\}$ and by β the number of times -1 occurs in this sequence so that $\alpha \geqq 0$, $\beta \geqq 0$ and $\alpha + \beta = +\infty$. The numbers α and β are uniquely determined by the matrix Q, i. e. they are independent of the special choice of the matrix O employed in transforming Q into the diagonal form $\| \epsilon_{pq} \|$. It will be necessary to distinguish the following four classes of matrices which are square roots of E:

$$
\begin{array}{lllll}
\text{I:} & \alpha = +\infty, & \beta = 2m & \text{for} & 0 \leqq m < +\infty, \\
\text{II:} & \alpha = +\infty, & \beta = 2m+1 & \text{for} & 0 \leqq m < +\infty, \\
\text{III:} & \alpha = m, & \beta = +\infty & \text{for} & 0 \leqq m < +\infty, \\
\text{IV:} & \alpha = +\infty, & \beta = +\infty. & &
\end{array} \tag{13}
$$

For the sake of brevity we shall denote an orthogonal matrix not containing -1 in its spectrum as a *Cayley matrix*. That not every orthogonal matrix is a Cayley matrix is trivial,‡ even for finite matrices; for example, no single reflection matrix (i. e. an orthogonal matrix of determinant -1) is a Cayley matrix. The justification for the name of this sub-manifold of the manifold of the orthogonal matrices is contained in the statement that the sub-manifold of the Cayley matrices may be put in a one-to-one correspondence with the manifold of the real bounded skew-symmetric matrices by means of the well known formula of Cayley,

$$O = (E+S)(E-S)^{-1} = (E-S)^{-1}(E+S). \tag{14}$$

† See "Hi", p. 145.

‡ Cf. M. Born and P. Jordan, "Elementare Quantenmechanik" (1930), p. 37, in which they point out that the representation

$$U = (E+iH)(E-iH)^{-1} = (E-iH)^{-1}(E+iH),$$

in which H is a bounded Hermitian matrix, is restricted to those unitary matrices U not containing -1 in their spectrum. The matrix U in this representation is then and only then a real matrix, i. e. an orthogonal matrix if $H = iS$, where S is a bounded real skew-symmetric matrix, the representation becoming, in this case, the well known representation of Cayley. For finite orthogonal matrices it is not true, as is sometimes stated, [cf. for instance H. W. Turnbull, "The Theory of Determinants, Matrices and Invariants" (1928), pp. 155-157] that every orthogonal matrix of determinant $+1$ permits the Cayley representation (cf. in particular p. 588). On the other hand, those orthogonal matrices of determinant $+1$, not permitting the Cayley representation, may be regarded as limiting cases of orthogonal matrices permitting the Cayley representation. In this connection cf. L. E. Dickson, "Modern Algebraic Theories" (1926), p. 102.

In order to demonstrate this let S in (14) be a bounded real skew-symmetric matrix so that $(E+S)^{-1}$ and $(E-S)^{-1}$ exist. The matrix O in (14) is then obviously real and is orthogonal for

$$OO' = O'O = (E+S)(E-S)^{-1}(E-S)(E+S)^{-1} = E.$$

Furthermore it is a Cayley matrix for, from (14),

$$O+E = (E+S)(E-S)^{-1} + E = (E+S)(E-S)^{-1} + (E-S)(E-S)^{-1} = (E+S+E-S)(E-S)^{-1} = 2(E-S)^{-1}$$

so that $(O+E)^{-1}$ exists, i. e. -1 is not a point in the spectrum of O. For the converse proposition, i. e. to any given Cayley matrix O there always exists a real bounded skew-symmetric matrix S for which (14) is valid, we consider the matrix

$$S = (O-E)(O+E)^{-1}, \tag{15}$$

which may be constructed for every Cayley matrix O. We first observe that because of

$$(O+E)(O-E) = (O-E)(O+E),$$

i. e.

$$(O-E)(O+E)^{-1} = (O+E)^{-1}(O-E),$$

we also have

$$S = (O+E)^{-1}(O-E). \tag{16}$$

Now the matrix S in (15), (16) is certainly real and bounded. In addition it is skew-symmetric, for

$$S' = (O'+E)^{-1}(O'-E) = (O'+E)^{-1}O'O(O'-E) = [O(O'+E)]^{-1}(E-O) = (O+E)^{-1}(E-O) = -S,$$

and may be employed as S in the representation (14) of O since

$$(E+S)(E-S)^{-1} = [E+(O-E)(O+E)^{-1}](O+E)(O+E)^{-1}[E-(O-E)(O+E)^{-1}]^{-1} = (O+E+O-E)(O+E-O+E)^{-1} = O.$$

That this correspondence set up between the manifold of the Cayley matrices and the bounded real skew-symmetric matrices is one to one is trivial. One may see, even for finite matrices, that the manifold of the Cayley matrices is only a very restricted subset of the manifold of all orthogonal matrices. For finite matrices the manifold of the Cayley matrices is only a sub-manifold of the manifold of the rotation matrices, not even every rotation matrix being a Cayley matrix. As an example of a matrix which is a rotation matrix and

not a Cayley matrix one has the matrix $-\boldsymbol{E}$ which for even n is a rotation matrix, but is for no single value of n a Cayley matrix. This fact is emphasized very forcibly by actually attempting the representation (14) for $\boldsymbol{O} = -\boldsymbol{E}$, thereby obtaining

$$-\boldsymbol{E} = (\boldsymbol{E} + \boldsymbol{S})(\boldsymbol{E} - \boldsymbol{S})^{-1},$$

and consequently

$$-\boldsymbol{E} + \boldsymbol{S} = \boldsymbol{E} + \boldsymbol{S} \quad \text{so that} \quad -\boldsymbol{E} = \boldsymbol{E},$$

an obvious contradiction. We are accordingly convinced that, although the Cayley representation (14) possesses no convergence difficulties, it is impossible to obtain, by its use, a parametric representation of the entire manifold of orthogonal matrices. In order to obtain a parametric representation of the entire manifold of the orthogonal matrices we employ the spectral theory which will yield a canonical exponential representation.

A Hermitian matrix is said to be *reduced* if its spectrum lies in the closed interval $[0, 2\pi]$ and if its point spectrum does not contain the point 2π. According to Wintner † a one to one correspondence is set up between the manifold of the unitary matrices and the manifold of the reduced Hermitian matrices by means of the exponential representation

$$\mathrm{U} = e^{i\mathrm{H}}, \tag{17}$$

in which U is a unitary matrix and H a reduced Hermitian matrix. The definition of the matrix e^{M} is of course

$$e^{\mathrm{M}} = \sum_{\nu=0}^{\infty} (1/\nu!)\mathrm{M}^{\nu},$$

which represents for every bounded matrix M a bounded matrix. Consequently if (M, N) is a commutable pair of bounded matrices, we have

$$e^{\mathrm{M}+\mathrm{N}} = e^{\mathrm{M}} \cdot e^{\mathrm{N}} = e^{\mathrm{N}} \cdot e^{\mathrm{M}}, \tag{18}$$

and if (M, N) is not only a commutable pair of matrices but also perpendicular [Cf. (7)]

$$e^{\mathrm{M}+\mathrm{N}} = e^{\mathrm{M}} + e^{\mathrm{N}} - \mathrm{E}. \tag{19}$$

For our purposes it is necessary to place the manifold of the bounded Hermitian matrices in a one-to-one correspondence with the combined manifolds

† See "W", pp. 268-277. For further literature cf. H. Weyl, *The Theory of Groups and Quantum Mechanics* (1931), p. 399.

of the bounded real symmetric and real skew-symmetric matrices by means of the linear representation

$$H = \pi B + iS, \quad B = \bar{B} = B', \quad S = \bar{S} = -S'. \tag{20}$$

In Theorem I we show that from a consideration of a trigonometric momentum problem with real "momenta" the entire manifold of the orthogonal matrices may, by means of the exponential representation (17), (20), i. e.

$$O = e^{iH} = e^{i[\pi B + iS]}, \tag{21}$$

be put in a one to one correspondence † with that sub-manifold of the reduced Hermitian matrices for which in (20)

$$B^2 = B, \qquad BS = SB, \tag{22}$$

i. e. B is a real symmetric idempotent (and therefore bounded) matrix and S a bounded real skew-symmetric matrix, commutable with B, and such that $\pi B + iS$ is a reduced Hermitian matrix.

On p. 606 we introduce the term *reduced skew-symmetric matrix* for those skew-symmetric matrices S, the entire spectrum of which is contained in the closed interval $[-i\pi, +i\pi]$ of the imaginary axis, and whose point spectrum does not contain the end-points $-i\pi$ and $+i\pi$. We then show in Lemma I that the manifold of the reduced Hermitian matrices (20), (22) may be put in a one-to-one correspondence with the manifold of the perpendicular matrix pairs $(E - B, S)$ in which S is a reduced skew-symmetric matrix. As a consequence of Lemma I we then obtain Theorem II which states that the entire manifold of the orthogonal matrices may, by means of the exponential representation (21), be put in a one-to-one correspondence with the manifold of the matrix pairs (B, S) where

$$\begin{gathered} B = \bar{B} = B', \; B^2 = B; \quad S = \bar{S} = -S'; \\ (E - B)S = S(E - B) = \| 0 \|, \quad S \text{ is reduced.} \end{gathered} \tag{23}$$

We then show in Theorem III that the one-to-one representation (21), (22) is in reality none other than the one-to-one representation

$$O = 2(E - B) - e^{-S} \tag{24}$$

in which the matrix pair (B, S) is identical with the matrix pair (B, S) in (21), (22).

In Lemma II on p. 615, we show that the manifold of the matrices Q which are square roots of E may be placed in a one-to-one correspondence

† See "W", p. 277.

with the manifold of the idempotent matrices I by means of the linear representation

$$Q = E - 2I. \tag{25}$$

An application of this lemma to Theorem I then yields Theorem IV on p. 615, in which the entire manifold of the orthogonal matrices is shown to be in a one-to-one correspondence, by means of the representation

$$O = Q \cdot e^{-S} = e^{-S} \cdot Q \tag{26}$$

with the manifold of the matrix pairs (Q, S) in which Q is a real symmetric square root of E (and therefore itself an orthogonal matrix) and S a bounded real skew-symmetric matrix, commutable with Q, and such that the Hermitian matrix $(\pi/2)(E - Q) + iS$ is reduced.

Since Theorem IV assigns to every orthogonal matrix O a uniquely determined real symmetric square root Q of E we may interpret the classification (13) of the real symmetric square roots of E as a classification of the orthogonal matrices into the four classes of p. 617. As immediate examples of the four types of orthogonal matrices we have the matrices Q of (124) themselves.

For normal forms, under orthogonal matrix transformations of the matrices of each of the four classes we obtain, as consequences of Theorems I and IV and Lemma I, the canonical forms given in (133) in which the matrices $Q^{(i)}$ and $S^{(i)}$ $(i = 1, 2, 3, 4)$ are defined in (124) and (125) respectively. The matrices $\boldsymbol{O}_{2m}$ and $\boldsymbol{O}_{2m+1}$ occurring in I and II of (133) are the canonical finite rotation matrices as defined in (131). From these canonical forms I and II of (133) it is clear at once that the orthogonal matrices of Classes I and II can be interpreted, with respect to a suitably orientated coördinate system, as rotations and reflections (in the geometrical sense of the word) in some subspace of the real Hilbert space, possessing a sufficiently high but finite number of dimensions, in which a subspace of infinitely many dimensions is not "affected" † by the transformation. The orthogonal matrices of Class III may accordingly be interpreted as representing those orthogonal transformations of the real Hilbert space which leave a subspace of at most a finite number of dimensions "unaffected" and the orthogonal matrices of Class IV as the orthogonal transformations of the real Hilbert space which

† For example we say that the transformation in the ordinary Euclidean 3-space possessing the matrix

$$\begin{Vmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{Vmatrix}$$

leaves a space of one dimension not "affected" by the transformation.

leave "unaffected" a subspace of infinitely many dimensions while simultaneously "affecting" a complement subspace of infinitely many dimensions. It may not be superfluous to remark that according to this classification the unit matrix E is an orthogonal matrix of Class I and the orthogonal matrix $-E$ of Class III. It is still an open question as to whether the orthogonal matrices of Classes III and IV may be divided into "rotations" and "reflections." In this connection cf. p. 592, where we are able to give a precise formulation of this unsolved problem.

During the last few years several authors have investigated rotations of the functional space, or what is the same, rotations in the Hilbert space. These papers, however, deal only with very special orthogonal matrices, supposing, for instance, that the determinant of the matrix is convergent and that the "deviation" of the orthogonal matrix from the unit matrix is a completely continuous (vollstetig in the sense of Hilbert) matrix. Even these very restrictive conditions are not sufficient for the validity of the determinant apparatus as employed by them. In this paper we deal with the totality of the orthogonal transformations of the Hilbert space and are, therefore, denied any recourse to the apparatus of determinants. For example the inaptness of the determinant apparatus is clearly revealed by attempting to differentiate, by its use, between the various classes (133) of the orthogonal matrices $Q^{(i)}$ of (124).

The Classes I and II of the orthogonal matrices O are then shown to be the natural extensions of the finite rotation and reflection matrices as defined on p. 579. The matrices of Classes III and IV obviously have no analogues for finite matrices. At this point there is introduced a definition which divides † the entire manifold of the infinite orthogonal matrices into two sub-manifolds in such a manner that the definition when applied to the manifold of the finite orthogonal matrices divides it into two sub-manifolds which are identical with the sub-manifolds of the finite rotation and reflection matrices as defined on p. 579. We define on p. 620 that sub-manifold of the orthogonal matrices, the elements O of which may be represented in the form e^G where G is a real bounded skew-symmetric matrix [not necessarily uniquely determined by O (Cf. p. 594)], i. e.

$$O = e^G, \qquad G = \bar{G} = -G', \tag{27}$$

as the manifold of the *rotation matrices* and the complement of this manifold with respect to the entire manifold of the orthogonal matrices [i. e. the manifold of those matrices not permitting the exponential representation (27)],

† Cf. p. 579.

as the manifold of the *reflection matrices.* This definition of the manifolds of the rotation and reflection matrices will be referred to as the exponential definition in order to distinguish it from the determinant definition on p. 579 which is utterly valueless for infinite orthogonal matrices since the determinant of an orthogonal matrix is convergent only for very special orthogonal matrices. In order to justify this definition we show on p. 620 in Lemma III that the exponential definition of the rotation and reflection matrices is, for finite matrices, identical with the determinant definition. In Theorem V on p. 621, we show that the matrices of Classes I and II are, in the sense of the exponential definition, respectively rotation and reflection matrices. The first part of Theorem V follows directly from the canonical form for the matrices of Class I. In order to prove the second part of Theorem V we show in Lemma IV on p. 622 that if θ (θ real) is a point in the spectrum of an arbitrary (not necessarily reduced) Hermitian matrix H then $e^{i\theta}$ is a point in the spectrum of $e^{i\mathrm{H}}$. As a corollary to Lemma IV we have on p. 624 that if $i\theta$ is a point in the spectrum of an arbitrary real bounded skew-symmetric matrix S then $e^{i\theta}$ is a point in the spectrum of e^{S} from which, and the canonical form II of (133), the second part of Theorem V is shown to follow.

The unsolved problem, mentioned on p. 591, as to whether the orthogonal matrices of Classes III and IV permit a resolution into "*rotations*" and "*reflections*" may now be put in a precise form as follows: Do the Classes III and IV contain both rotation and reflection matrices as defined in (27)? That there exist rotation matrices in the sense of the exponential definition (27) is trivial since the orthogonal matrices $\mathrm{Q}^{(3)}$ and $\mathrm{Q}^{(4)}$ of (124) are of Classes III and IV respectively and obviously permit the exponential representation (27). The unsolved part of the problem then requires the demonstration of the existence or the non-existence of reflection matrices, as defined in (27), for each of the Classes III and IV. Put more directly, do the matrices of the Classes III and IV all permit the exponential representation (27)? Since the property of a matrix to be a rotation matrix, in the sense of the exponential definition (27), is obviously invariant † with respect to orthogonal transformations, it is sufficient to treat this question for the canonical forms III and IV of (133) for the matrices of Classes III and IV. From the fact that the orthogonal matrices $\mathrm{Q}^{(3)}$ and $\mathrm{Q}^{(4)}$ in the canonical forms (133) are both rotation matrices, it is trivial that if the

† This invariance property is trivial inasmuch as follows from the definition of e^{G} on p. 13 that if O is an arbitrary orthogonal matrix then $\mathrm{O}e^{\mathrm{G}}\mathrm{O}^{-1} = e^{\mathrm{OGO}^{-1}}$ in which OGO^{-1} is again a bounded real skew-symmetric matrix.

rotation matrices [i. e. those orthogonal matrices permitting the representation (27)] form a group, the canonical matrices III and IV of (133) are necessarily rotation matrices [i. e. also permit the representation (27)]. It may not be superfluous to point out that, because of the special character of the matrices occurring in the canonical matrices III and IV of (133), the condition that the rotation matrices (27) form a group is not a necessary condition for the matrices of Classes III and IV to be rotation matrices (27). A treatment of the above unsolved problems seem to require a detailed investigation of orthogonal and real skew-symmetric matrices possessing continuous spectra. Connected with this problem is the question † as to whether the manifold of all orthogonal matrices forms a connected manifold or not (with respect to some natural metric in the manifold of all matrices). For the orthogonal transformations of a space with a finite number of dimensions the manifold of all orthogonal matrices is not connected (with respect to every natural metric of the manifold) but is divided into two separate manifolds (Cf. p. 579). In other words the Euclidean space of a finite number of dimensions may be said to be an "orientable" space. It has not been possible as yet to answer the question, on the basis of the parametrical representation of orthogonal matrices given in the paper, as to whether the real Hilbert space forms, in this sense, an orientable space or not. In order that the manifold of all orthogonal matrices be connected it is necessary that each of the sub-manifolds in (133) be connected. It seems likely that for the answer to the above question the extensive use of the theory of the spectral matrix of the orthogonal matrices is unavoidable, the problems treated in the present paper depending upon the theory of continuous spectra of unitary matrices (cf. "W", pp. 257-258).

Theorems I, II, III and IV all give a complete parametric representation of the manifold of the orthogonal matrices and we saw that (27) where G is an arbitrary real bounded skew-symmetric matrix, may be defined as the parametric representation of the manifold of the rotation matrices of the Hilbert space. This parametric representation of the rotation matrices is the substitute † in Hilbert space for the Euler-Hurwitz factorization ‡ into products of "*two-block*" rotation matrices which breaks down for infinite

† See "W", p. 277.

‡ L. Euler, *Commentationes arithmeticae collectae*, Vol. 1, p. 427 and A. Hurwitz, "Ueber die Erzeugung der Invarianten durch Integration," *Göttingen Nachrichten* (1897), pp. 76-77. In connection with the work of Hurwitz, cf. I. Schur, "Neue Anwendungen der Integralrechnung auf Probleme der Invariantentheorie," *Sitzungsberichte der Preussischen Akademie der Wissenschaften* (1924), p. 196, where a statement of Hurwitz is corrected.

matrices since there exists no last dimension. There are even other obstacles to a generalization of the Euler-Hurwitz representation inasmuch as the following example shows that a matrix which is (even in the sense of a "*weak*" convergence) the limit of a sequence of orthogonal "*two block*" matrices is not necessarily itself an orthogonal matrix. Denote by E_{pq} the "*two- block*" orthogonal matrix obtained from the infinite unit matrix E by interchanging the p-th and q-th columns (or what amounts to the same thing, by interchanging the p-th and q-th rows). Consider the sequence of orthogonal matrices $\{O_i\}$ defined by $O_1 = E$, $O_i = O_{i-1}E_{i-1\,i}$ $(i = 2, 3, \cdots)$, i. e. the sequence

$$O_1 = E, \quad O_2 = E_{12}, \quad O_3 = E_{12}E_{23}, \cdots,$$

for which

$$\lim_{n\to+\infty} O_n = \begin{Vmatrix} 0 & 0 & 0 & 0 & . & . & . \\ 1 & 0 & 0 & 0 & . & . & . \\ 0 & 1 & 0 & 0 & . & . & . \\ 0 & 0 & 1 & 0 & . & . & . \\ . & . & . & . & & & \\ . & . & . & . & & & \\ . & . & . & . & & & \end{Vmatrix},$$

and which is certainly not an orthogonal matrix.

It may be mentioned that to a given rotation matrix O there may exist infinitely many real bounded skew-symmetric matrices which may be employed as the matrix G in (27). This is illustrated, even for finite matrices, by the example

$$\begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix} = \exp \begin{Vmatrix} 0 & (\theta + 2k\pi) \\ -(\theta + 2k\pi) & 0 \end{Vmatrix}; \quad (k = 0, \pm 1, \cdots).$$

It would be desirable to obtain a simple sub-manifold of the matrices G in (27) so that the matrices of this sub-manifold are in one-to-one correspondence with the rotation matrices. Certain considerations for finite matrices seem to indicate that this may be achieved for finite matrices by restricting the representation (27) to a sub-manifold analogous to the manifold of the reduced skew-symmetric matrices, and it is hoped to treat this question among others in a later appearing paper on the finite matrices.

For finite matrices it is possible † to obtain, for $n = 2$ and $n = 3$, the formulas of Rodrigues from the exponential representation (27).

---

† Cf. G. Galanti, "Algoritmi di calcolo motoriale," *Atti Accademia Nazionale dei Lincei*, Vol. 6, series 13 (1931), pp. 861-866.

Finally, it may be mentioned that the methods and results of the present paper hold, precisely as the spectral theory of unitary transformations or the method of trigonometric momentum problem used in this paper, not only for matrices but also for operators in the Hilbert space. However, since the possibility of such generalizations is known to be trivial we prefer, for clearness, to deal with matrices only.

I. Reality Discussion of the Trigonometrical Moment Problem of Herglotz.

We consider the trigonometrical moment problem †

$$2\pi c_n = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi)\,; \qquad (n = 0, \pm 1, \cdots), \tag{28}$$

together with its homogeneous special case

$$0 = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi)\,; \qquad (n = 0, \pm 1, \cdots), \tag{28'}$$

in which the left-hand members are preassigned real or complex numbers and $\tau(\phi)$ an unknown real or complex function of the real argument ϕ whose definition is desired in the closed interval $0 \leqq \phi \leqq 2\pi$. In the following we shall consider only real solutions $\tau(\phi)$ and therefore suppose that

$$c_n = \bar{c}_{-n}; \qquad (n = 0, 1, \cdots). \tag{29}$$

It will be convenient to employ the notations

$$\begin{aligned} \underset{\beta-0}{\Delta} f(\phi) &= f(\beta) - f(\beta - 0), \\ \underset{\beta}{\Delta} f(\phi) &= f(\beta + 0) - f(\beta - 0), \\ \underset{\beta+0}{\Delta} f(\phi) &= f(\beta + 0) - f(\beta), \end{aligned} \tag{30}$$

(provided these limits exist) for the various "springs" of an arbitrary real or complex function $f = f(\phi)$ at the point $\phi = \beta$ and the customary notations $[a, b]$ and (a, b) for the closed and open intervals $a \leqq \phi \leqq b$ and $a < \phi < b$ respectively.

We suppose that the moment problem (28) permits a solution $\tau(\phi)$ of bounded variation in the interval of its definition $[0, 2\pi]$ so that the limits

† G. Herglotz, "Über Potenzreihen mit positiven, reellen Teil im Einheitskreis," *Königlich Sächsische Gesellschaft der Wissenschaften zu Leipsig, Mathematisch-Physische Klasse* (1911), p. 501-511. This paper will be referred to as "He".

(31) $\tau(\phi-0)$ for $0<\phi\leqq 2\pi$; $\tau(\phi+0)$ for $0\leqq\phi<2\pi$,

exist. It is known † from the general theory of Stieltjes integrals that if (28) possesses a solution $\tau(\phi)$ which is of bounded variation in the interval of its definition $[0, 2\pi]$ it is uniquely determined, up to an additive constant, on the set of its continuity points contained in the interval $(0, 2\pi)$, its values at its inner discontinuity points (i. e. for $0<\phi<2\pi$) being assignable at will. Since the set of continuity points of $\tau(\phi)$ in the interval $(0, 2\pi)$ is necessarily everywhere dense in this interval the values of $\tau(0+0)$ and $\tau(2\pi-0)$ may then be calculated [cf. (31)] from the values of $\tau(\phi)$ at its continuity points and the values of $\tau(0)$ and $\tau(2\pi)$ chosen at will, provided that the conditions

$$(32)\quad \underset{0+0}{\Delta}\,\tau(\phi)+\underset{2\pi-0}{\Delta}\,\tau(\phi)=2\pi c_0-\int_{0+0}^{2\pi-0}d\tau(\phi)=2\pi c_n-\int_{0+0}^{2\pi-0}e^{in\phi}\,d\tau(\phi);\qquad (n=0,\pm 1,\cdots),$$

which arise from (28), are satisfied.‡ The only solutions of bounded variation of the homogeneous moment problem (28′) may in turn be completely characterized by the properties §

$$(33)\quad \tau(\phi)=\alpha=\bar{\alpha}=\text{const. at the continuity points of } \tau(\phi) \text{ in } (0, 2\pi),$$

$$(34)\quad \underset{0+0}{\Delta}\,\tau(\phi)+\underset{2\pi-0}{\Delta}\,\tau(\phi)=0,$$

inasmuch as, in (32), we have $c_0=0$ and $\int_{0+0}^{2\pi-0}d\tau(\phi)=0$.

A further possible specialization of the character of the solution $\tau(\phi)$ of (28) is that it be monotone non-decreasing. The necessary and sufficient conditions on the c_n in order that the moment problem (28) possess a monotone non-decreasing solution $\tau(\phi)$ are given by Herglotz ¶ and state that the truncated matrices (Abschnittsmatrizen) of the infinite Hermitian matrix $\| c_{\mu-\nu} \|$, introduced by Toeplitz, be not negative definite, i. e.

$$(35)\quad \sum_{\mu=1}^{n}\sum_{\nu=1}^{n}c_{\mu-\nu}\xi_\mu\bar{\xi}_\nu\geqq 0 \text{ for arbitrary } \xi_1, \xi_2, \cdots; \qquad (n=1, 2, \cdots),$$

where $c_{\mu-\nu}=c_k=\bar{c}_{-k}$ for $k=\mu-\nu$.

---

† Cf. F. Hausdorff, "Über das Momentenproblem für ein endliches Interval," *Mathematische Zeitschrift*, Vol. 16 (1923), pp. 220-248. This paper will be referred to as "Ha". Cf. also O. Perron, "Die Lehre von den Kettenbrüchen" (1913), pp. 362-374. This book will be referred to as "P".

‡ See "W", p. 268.

§ See "Ha", p. 241.

¶ See "He".

We denote by $\mathrm{E} = \| \delta_{pq} \|$ the unit matrix and by $\Phi[\mathrm{A}; x]$ the compound form (Kopplungsform)

$$\Phi[\mathrm{A}; x] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p \bar{x}_q, \tag{36}$$

of an arbitrary bounded matrix $\mathrm{A} = \| a_{pq} \|$. With these notations we consider the trigonometric moment problem † arising from (28) if one places

$$c_n = \Phi[\mathrm{U}^n; x]; \qquad (n = 0, \pm 1, \cdots), \tag{37}$$

i. e., the moment problem

$$2\pi\Phi[\mathrm{U}^n; x] = \int_0^{2\pi} e^{in\phi}\, d\tau(\phi; x), \qquad |x| = 1; \ (n = 0, \pm 1, \cdots), \tag{38}$$

where U is a given unitary matrix so that for every unit vector x [i. e., $|x| = 1$] there belongs the sequence of momenta c_k defined in (37) satisfying the conditions (29). It is known that this moment problem always possesses a real solution $\tau(\phi; x)$ which depends upon the vector x and is, as a matter of fact, a bounded Hermitian form

$$\tau(\phi; x) \equiv \Phi[\| \tau_{pq}(\phi) \|; x], \tag{39}$$

where

$$\tau_{pq}(\phi) = \bar{\tau}_{qp}(\phi), \qquad 0 \leqq \phi \leqq 2\pi,$$

and the indices p, q take, as will always be understood in the remainder of the paper, unless explicitly stated to the contrary, independently the values $1, 2, \cdots$. The Hermitian form (39) is then, for an arbitrary given unit vector x, uniquely determined up to an additive constant at all of its continuity points in the interval $(0, 2\pi)$. Moreover, since the conditions (35) are known to be fulfilled by the momenta (37), for every unit vector x, the Hermitian form (39) may be chosen to be, for an arbitrary given unit vector x, a monotone non-decreasing function of ϕ in the interval $(0, 2\pi)$. The values $\tau(0, x)$ and $\tau(2\pi, x)$ of the Hermitian form at the end points may then, for an arbitrary given unit vector x, be assigned at will subject to the condition that $\tau(\phi; x)$ be monotone in the closed interval $[0, 2\pi]$ and that

$$\underset{0+0}{\Delta}\, \tau(\phi; x) + \underset{2\pi-0}{\Delta}\, \tau(\phi; x) = 2\pi - \int_{0+0}^{2\pi-0} d\tau(\phi; x). \tag{40}$$

We now remove the arbitrary character of the solution $\tau(\phi; x)$ by prescribing, for an arbitrary given unit vector x, the following normalizing conditions:

† See "W", p. 269.

$$\tau(0;x) \equiv 0, \tag{41}$$

$$\tau(\phi - 0;x) \equiv \tau(\phi;x) \quad \text{for} \quad 0 < \phi < 2\pi, \tag{42}$$

$$\underset{2\pi-0}{\Delta}\, \tau(\phi;x) \equiv 0. \tag{43}$$

The normalizing condition (42) serves to fix the value of the solution $\tau(\phi;x)$ at its discontinuity points in the interval $(0, 2\pi)$. The prescribing of the normalizing condition (43) then from (40) serves to determine uniquely the "spring" $\underset{0+0}{\Delta}\, \tau(\phi;x)$. Finally the normalizing condition (41) fixes the value of the additive constant so that the solution $\tau(\phi;x)$ is uniquely determined at all points of the closed interval $[0, 2\pi]$ since, for $\phi = 2\pi$, we have from (41) and (38), for $n = 0$,

$$\tau(2\pi;x) \equiv 2\pi, \qquad |x| = 1. \tag{44}$$

The matrix of the Hermitian form under the normalizing conditions (41), (42) and (43) has been introduced by Wintner † and designated by him as the spectral matrix of U. Since (38), (41), (42), (43) and (44) hold for an arbitrary unit vector x the elements of the spectral matrix $\| \tau_{pq}(\phi) \|$ obviously possess the following properties ‡:

$$2\pi U^n = \| \int_0^{2\pi} e^{in\phi}\, d\tau_{pq}(\phi) \|; \qquad (n = 0, \pm 1, \cdots), \tag{38'}$$

$$\tau_{pq}(0) = 0, \tag{41'}$$

$$\tau_{pq}(\phi - 0) = \tau_{pq}(\phi) \quad \text{for} \quad 0 < \phi < 2\pi, \tag{42'}$$

$$\underset{2\pi-0}{\Delta}\, \tau_{pq}(\phi) = 0 \quad \text{i. e.} \quad \tau_{pq}(2\pi - 0) = \tau_{pq}(2\pi), \tag{43'}$$

$$\tau_{pq}(2\pi) = 2\pi\delta_{pq}. \tag{44'}$$

If now the unitary matrix is assumed to be real, i. e. U is an orthogonal matrix $U = \bar{U} = O$, the elements of the spectral matrix possess in addition to the above mentioned properties the following two properties:

I. *In order that the elements $\tau_{pq}(\phi)$ of the spectral matrix of the unitary matrix* U *be the elements of the spectral matrix of a real unitary matrix* $U = \bar{U}$, *i. e. an orthogonal matrix* O, *it is necessary and sufficient that*

$$\omega_{pq}(\phi) = b\delta_{pq} \tag{45}$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$, where

† See "W", p. 269.

‡ Cf. the uniqueness property of compound forms mentioned on p. 580.

$$\omega_{pq}(\phi) = \bar{\tau}_{pq}(\phi) + \tau_{pq}(2\pi - \phi) \quad \textit{and} \quad b = \bar{b} = \textit{const.} \tag{46}$$

II. *The "springs"* $\underset{0+0}{\Delta}\, \tau_{pq}(\phi)$ *of the elements of the spectral matrix of an orthogonal matrix* O *are real.*

The demonstration of the two statements I and II proceeds simultaneously and we begin by showing the necessity of (45). We accordingly assume $U = \bar{U} = O$ so that the conjugate equations of (38′) become

$$2\pi O^n = \| \int_0^{2\pi} e^{-in\phi}\, d\bar{\tau}_{pq}(\phi) \|; \qquad (n = 0, \pm 1, \cdots). \tag{47}$$

On the other hand if we introduce a new integration variable θ in (38′) by means of the transformation

$$\phi = 2\pi - \theta, \tag{48}$$

we obtain

$$-2\pi O^n = -\| \int_{2\pi}^{0} e^{in(2\pi-\theta)}\, d\tau_{pq}(2\pi - \theta) \| = \| \int_0^{2\pi} e^{-in\theta}\, d\tau_{pq}(2\pi - \theta) \|. \tag{49}$$

$$(n = 0, \pm 1, \cdots).$$

On adding (47) and (49) we have from (46)

$$\| 0 \| = \| \int_0^{2\pi} e^{-in\phi}\, d\omega_{pq}(\phi) \|; \qquad (n = 0, \pm 1, \cdots), \tag{50}$$

so that the Hermitian form $\omega(\phi; x) = \Phi[\| \omega_{pq}(\phi) \|; x]$ is, for any arbitrary given unit vector x, a solution of the homogeneous moment problem (28′). Furthermore the Hermitian form $\Phi[\| \tau_{pq}(\phi) \|; x]$ and consequently the Hermitian form $\Phi[\| \bar{\tau}_{pq}(\phi) \|; x]$ are, independently of the unit vector x, monotone non-decreasing functions of ϕ in the interval $[0, 2\pi]$. The Hermitian form $\Phi[\| \tau_{pq}(2\pi - \phi) \|; x]$ is, on the contrary, monotone non-increasing in $[0, 2\pi]$, for any given unit vector x. Accordingly the Hermitian form $\omega(\phi; x)$ is of bounded variation [cf. (46)] in the interval $[0, 2\pi]$ for any arbitrary given unit vector x. From the uniqueness property (33) of solutions of bounded variation of the homogeneous moment problem (28′) we have, for an arbitrary given unit vector x,

$$\omega(\phi; x) \equiv \Phi[\| \bar{\tau}_{pq}(\phi) + \tau_{pq}(2\pi - \phi) \|; x] = b = \text{const.}, \tag{51}$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$, and consequently, since (51) holds for an arbitrary unit vector x, the necessity of the condition (45) of I. Furthermore, since $\| \tau_{pq}(\phi) \|$ is Hermitian, the constant b in (51) is real.

The statement II is an immediate consequence of the necessary condition

(45) and the normalizing conditions (41′), (43′), (44′). From (45) we obtain on observing (41′)

$$\bar{\tau}_{pq}(\phi) - \bar{\tau}_{pq}(0) = b\delta_{pq} - \tau_{pq}(2\pi - \phi) \tag{52}$$

at the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$. Since the continuity points of $\omega_{pq}(\phi)$, which is obviously of bounded variation in $[0, 2\pi]$, are certainly everywhere dense in the interval $(0, 2\pi)$ we have for $\phi \to +0$ [cf. (30)] from (43′) and (44′)

$$\underset{0+0}{\Delta}\, \bar{\tau}_{pq}(\phi) = (b - 2\pi)\delta_{pq} \quad \text{where} \quad b = \bar{b} = \text{const.}, \tag{53}$$

i. e. II is proven.

We now show the sufficiency of the condition (45) in I. From (38′) for $n = 1$ we have as a consequence of (43′)

$$2\pi U = \| \int_0^{2\pi} e^{i\phi}\, d\tau_{pq}(\phi) \| = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \| - \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \|,$$

which may, on introducing a new integration variable θ by means of (48) in $\int_{\pi+0}^{2\pi-0}$, be written in the form

$$2\pi U = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} e^{i\phi}\, d\tau_{pq}(\phi) \| - \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \| - \| \int_{0+0}^{\pi-0} e^{-i\theta}\, d\tau_{pq}(2\pi - \theta) \|. \tag{54}$$

We now utilize (45) to obtain

$$\int_{0+0}^{\pi-0} e^{-i\phi}\, d\omega_{pq}(\phi) = \int_{0+0}^{\pi-0} e^{-i\phi}\, d\bar{\tau}_{pq}(\phi) + \int_{0+0}^{\pi-0} e^{-i\phi}\, d\tau_{pq}(2\pi - \phi) = 0, \tag{55}$$

inasmuch as the values of $\omega_{pq}(\phi)$ at the inner discontinuity points are immaterial in the Stieltjes integration. From (55) we may write (54) in the form

$$2\pi U = \| \underset{0+0}{\Delta}\, \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} e^{i\phi} d\tau_{pq}(\phi) + \int_{0+0}^{\pi-0} e^{-i\phi} d\bar{\tau}_{pq}(\phi) \| - \| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \|,$$

so that from II there only remains to show that $\| \underset{\pi}{\Delta}\, \tau_{pq}(\phi) \|$ is a real matrix. This follows immediately on writing (45) in the form

$$\tau_{pq}(2\pi - \phi) - \tau_{pq}(\phi) = b\delta_{pq} - [\tau_{pq}(\phi) + \bar{\tau}_{pq}(\phi)],$$

valid for the continuity points of $\omega_{pq}(\phi)$ in $(0, 2\pi)$, and then taking the limit for $\phi \to \pi - 0$.

II. Application of the Trigonometrical Moment Problem to Orthogonal Matrices.

Among the properties of the spectral matrix $\| \tau_{pq}(\phi) \|$ of a given unitary matrix U, Wintner † has shown that it is possible to employ it in the construction of a bounded Hermitian matrix H

$$\mathrm{H} = \| (1/2\pi) \int_0^{2\pi} \phi d\tau_{pq}(\phi) \|, \tag{56}$$

the elements of the spectral matrix $\| \sigma_{pq}(\phi) \|$ of which are given by

$$\begin{aligned} \sigma_{pq}(\phi) &= 0 && \text{for} \quad -\infty < \phi \leqq 0, \\ 2\pi\sigma_{pq}(\phi) &= \tau_{pq}(\phi) && \text{for} \quad 0 \leqq \phi \leqq 2\pi, \\ \sigma_{pq}(\phi) &= \delta_{pq} && \text{for} \quad 2\pi \leqq \phi < +\infty. \end{aligned} \tag{57}$$

Conversely if H is a bounded Hermitian matrix whose spectral matrix $\| \sigma_{pq}(\phi) \|$ fulfills the conditions

$$\begin{aligned} \sigma_{pq}(\phi) &= 0 && \text{for} \quad -\infty < \phi \leqq 0, \\ \sigma_{pq}(2\pi - 0) &= \sigma_{pq}(\phi) = \delta_{pq} && \text{for} \quad 2\pi \leqq \phi < +\infty, \end{aligned} \tag{58}$$

(i. e. $\phi = 2\pi$ is not a point of the point spectrum of H) and if a matrix $\| \tau_{pq}(\phi) \|$ be introduced by means of the definitions

$$\tau_{pq}(\phi) = 2\pi\sigma_{pq}(\phi) \quad \text{for} \quad 0 \leqq \phi \leqq 2\pi, \tag{59}$$

then the matrix

$$\mathrm{U} = e^{i\mathrm{H}} = \| (1/2\pi) \int_0^{2\pi} e^{i\phi} \, d\tau_{pq}(\phi) \|, \tag{60}$$

is a unitary matrix possessing as its spectral matrix the matrix $\| \tau_{pq}(\phi) \|$. If we designate a Hermitian matrix whose spectrum lies in the interval $[0, 2\pi]$ and whose point spectrum does not contain the point 2π, as a reduced Hermitian matrix we may sum up the above results in the following form: There exists between the unitary matrices and the reduced Hermitian matrices a one-to-one correspondence, i. e. to a given unitary matrix U there is associated a reduced Hermitian matrix H by means of (56) and to this same reduced Hermitian matrix H there is associated, by means of (60), the previously given unitary matrix U.

In the following we propose to apply the results of the previous section in order to characterize that sub-class of the reduced Hermitian matrices which is in one-to-one correspondence with the real unitary matrices, i. e.

† See "W", pp. 268-277.

the orthogonal matrices. In any case we may write any Hermitian matrix uniquely in the form

$$H = \pi B + iS, \tag{61}$$

where B is a real symmetric matrix and S a real skew-symmetric matrix so that our problem † is to characterize the matrix pair (B, S) for which the reduced Hermitian matrix $H = \pi B + iS$ has the property that the matrix $U = e^{iH}$ is real. As an answer to this problem we show

THEOREM I. *The necessary and sufficient conditions in order that the reduced Hermitian matrix* $H = \pi B + iS$ *yield a real unitary matrix* O, *i. e. an orthogonal matrix, in the one-to-one representation*

$$O = e^{iH} = e^{i(\pi B + iS)}, \tag{62}$$

are that S *be a bounded real skew-symmetric matrix and* B *a real symmetric idempotent matrix which is commutable with* S, *i. e.*

$$B = \bar{B} = B', \quad B^2 = B; \qquad S = \bar{S} = -S', \tag{63a}$$

$$BS = SB. \tag{63b}$$

Before beginning the demonstration of Theorem I we note that the representation (62) may, from (63b) and the distributive property (18) of the exponential with respect to commutable matrices, be given the form

$$O = e^{i\pi B} \cdot e^{-S} = e^{-S} \cdot e^{i\pi B}. \tag{62'}$$

It may also be mentioned that the representation (62) or its equivalent (62′) is "covariant" under orthogonal matrix transformation. More precisely if V is any orthogonal matrix then we have from (62)

$$VOV^{-1} = \exp\,[i\pi(VBV^{-1} + iVSV^{-1})],$$

where VBV^{-1} is a real symmetric idempotent matrix [cf. p. 585] commutable with the bounded real skew-symmetric matrix VSV^{-1} and $\pi VBV^{-1} + iVSV^{-1}$ is a reduced Hermitian matrix if $\pi B + iS$ is a reduced Hermitian matrix. This "covariance" property is important in what follows inasmuch as without it the classification of the orthogonal matrices in four classes as developed below would be meaningless.

We begin with a demonstration of the necessity of the conditions (63a) and (63b) and shall retain the notation $\| \tau_{pq}(\phi) \|$ for the spectral matrix of a unitary matrix U. First we show that,

† See "W", p. 277.

$$(64) \qquad \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) = \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi),$$

$$(65) \qquad \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) = 2\pi \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) - \int_{0+0}^{\pi-0} \phi d\bar{\tau}_{pq}(\phi).$$

On introducing a new integration variable θ by means of (48) we obtain for the left-hand members of (64) and (65) respectively

$$(66) \qquad \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) = \int_{2\pi-0}^{\pi+0} d\bar{\tau}_{pq}(2\pi - \theta).$$

$$(67) \qquad \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) = \int_{\pi-0}^{0+0} (2\pi - \theta) d\tau_{pq}(2\pi - \theta).$$

From (45) we obtain

$$(68) \qquad \int_{2\pi-0}^{\pi+0} d\omega_{pq}(\phi) = \int_{2\pi-0}^{\pi+0} d\bar{\tau}_{pq}(\phi) + \int_{2\pi-0}^{\pi+0} d\tau_{pq}(2\pi - \phi) = 0,$$

$$(69) \qquad \int_{\pi-0}^{0+0} (2\pi - \phi) d\omega_{pq}(\phi) = \int_{\pi-0}^{0+0} (2\pi - \phi) d\bar{\tau}_{pq}(\phi) + \int_{\pi-0}^{0+0} (2\pi - \phi) d\tau_{pq}(2\pi - \phi) = 0,$$

inasmuch as the values of $\omega_{pq}(\phi)$ at its inner discontinuity points play no rôle in the Stieltjes integration in (68) and (69). Formulas (64) and (65) obviously follow from (66) and (67) by virtue of (68) and (69), respectively. The reduced Hermitian matrix H in the exponential representation of Theorem I may always be calculated from the spectral matrix of the given unitary matrix U by means of (56). For the treatment of a real unitary matrix $U = \bar{U} = O$ we express the Stieltjes integration in (56) as follows:

$$(70) \qquad 2\pi H = \| \int_{0+0}^{\pi-0} \phi d\tau_{pq}(\phi) \| + \pi \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} \phi d\tau_{pq}(\phi) \|,$$

inasmuch as [cf. (43′) on p. 598]

$$\int_{0}^{0+0} \phi d\tau_{pq}(\phi) = \int_{2\pi-0}^{2\pi} \phi d\tau_{pq}(\phi) = 0.$$

By means of (65) we may rewrite (70) in the form

$$2\pi H = \| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \| + \pi \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + 2\pi \| \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) \|,$$

which in turn may be replaced by

$$(71)\quad 2\pi H = \| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \| + \pi \| \int_{0+0}^{\pi-0} d[\bar{\tau}_{pq}(\phi) - \tau_{pq}(\phi)] \| + \pi \| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \| + \pi \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + \pi \| \int_{0+0}^{\pi-0} d\bar{\tau}_{pq}(\phi) \|.$$

We introduce a bounded real skew-symmetric matrix S by means of

$$(72)\quad 2\pi S = -i \| \int_{0+0}^{\pi-0} \phi d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \| - i\pi \| \int_{0+0}^{\pi-0} d[\bar{\tau}_{pq}(\phi) - \tau_{pq}(\phi)] \| = -i \| \int_{0+0}^{\pi-0} (\phi - \pi) d[\tau_{pq}(\phi) - \bar{\tau}_{pq}(\phi)] \|,$$

and employ it, together with (64), to abbreviate (71) to

$$(73)\quad 2H = 2iS + \| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \| + \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi) \|.$$

Now from (38′) [cf. p. 598] for $n = 0$ we have

$$2\pi E = \| \int_{0}^{2\pi} d\tau_{pq}(\phi) \| = \| \underset{0+0}{\Delta} \tau_{pq}(\phi) \| + \| \int_{0+0}^{\pi-0} d\tau_{pq}(\phi) \| + \| \underset{\pi}{\Delta} \tau_{pq}(\phi) \| + \| \int_{\pi+0}^{2\pi-0} d\tau_{pq}(\phi) \|,$$

inasmuch as from (43′)

$$\int_{2\pi-0}^{2\pi} d\tau_{pq}(\phi) = 0,$$

so that (73) may be further shortened to

$$(74)\quad H = iS + \pi E - (1/2) \| \underset{0+0}{\Delta} \tau_{pq}(\phi) \| = iS + \pi [E - (1/2\pi) \| \underset{0+0}{\Delta} \tau_{pq}(\phi) \|],$$

in which S is the bounded real skew-symmetric matrix introduced in (72) and $\| \underset{0+0}{\Delta} \tau_{pq}(\phi) \|$, since $\| \tau_{pq}(\phi) \|$ is a bounded Hermitian matrix, is, from II on p. 599, a bounded real symmetric matrix. We introduce a bounded real [cf. (61)] symmetric matrix B by means of

$$B = E - (1/2\pi) \| \underset{0+0}{\Delta} \tau_{pq}(\phi) \|,$$

and which, from (59), may obviously be written in the form

$$(75)\quad B = E - \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|,$$

so that, inasmuch as $\| \sigma_{pq}(\phi) \|$ is the spectral matrix of the reduced Hermitian matrix H we have †

† See "Hi", pp. 141-145. Cf. also "S", p. 160 and p. 174.

$$B^2 = E - 2 \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| + \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|^2 = E - \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| = B,$$

i. e. B is a real symmetric idempotent matrix so that the necessity of (63a) is demonstrated.

In order to demonstrate the necessity of (63b) we write (72) in the form

$$2\pi S = \lim_{n \to +\infty} 2\pi S_n, \tag{76}$$

where

$$2\pi S_n = -i \Big\| \sum_{\nu=1}^{n-2} \left(\frac{\nu}{n}\pi - \pi\right) \Big\{ \left[\tau_{pq}\left(\frac{\nu+1}{n}\pi\right) - \tau_{pq}\left(\frac{\nu}{n}\pi\right)\right] - \left[\bar{\tau}_{pq}\left(\frac{\nu+1}{n}\pi\right) - \bar{\tau}_{pq}\left(\frac{\nu}{n}\pi\right)\right] \Big\} \Big\|.$$

or, according to (59),

$$S_n = -i \Big\| \sum_{\nu=1}^{n-2} \left(\frac{\nu}{n}\pi - \pi\right) \Big\{ \left[\sigma_{pq}\left(\frac{\nu+1}{n}\pi\right) - \sigma_{pq}\left(\frac{\nu}{n}\pi\right)\right] - \left[\bar{\sigma}_{pq}\left(\frac{\nu+1}{n}\pi\right) - \bar{\sigma}_{pq}\left(\frac{\nu}{n}\pi\right)\right] \Big\} \Big\|. \tag{77}$$

We now apply the well-known theorem of Hilbert † that

$$\| \sigma_{pq}(\phi_1) \| \ \| \sigma_{pq}(\phi_2) \| = \| \sigma_{pq}(\phi_2) \| \ \| \sigma_{pq}(\phi_1) \|,$$

to obtain

$$\| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| \ \| \sigma_{pq}(\nu\pi/n) \| = \| \sigma_{pq}(\nu\pi/n) \| \ \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|; \tag{78}$$

$$(\nu = 1, 2, \cdots, n-1); \ (n = 2, 3, \cdots),$$

and hence also

$$\| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \| \ \| \bar{\sigma}_{pq}(\nu\pi/n) \| = \| \bar{\sigma}_{pq}(\nu\pi/n) \| \ \| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|; \tag{79}$$

$$(\nu = 1, 2, \cdots, n-1); \ (n = 2, 3, \cdots),$$

inasmuch as, from II on p. 599, the matrix $\| \underset{0+0}{\Delta} \sigma_{pq}(\phi) \|$ is a real matrix.

From (75), (77), (78) and (79) we accordingly have

$$(E - B)S_n = S_n(E - B),$$

from which and (76) then follows, for $n \to +\infty$, the necessity of (63b).

We now demonstrate the converse part of Theorem I, i. e. we assume that $H = \pi B + iS$ is a reduced Hermitian matrix for which the matrices

† See "Hi", pp. 141-145.

B and S fulfill the conditions (63a) and (63b) and demonstrate that U is a real matrix. We precede the proof by the derivation of the identity

$$e^{i\theta \mathrm{I}} = \mathrm{E} + (e^{i\theta} - 1)\mathrm{I}, \tag{80}$$

where θ is an arbitrary real or complex number and I any bounded idempotent matrix (cf. p. 585). Since I is assumed to be bounded we have the convergent development (cf. p. 588)

$$e^{i\theta \mathrm{I}} = \sum_{\nu=0}^{+\infty} \frac{(i\theta \mathrm{I})^{\nu}}{\nu!} = \sum_{\nu=0}^{+\infty} \frac{(i\theta \mathrm{I})^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{(i\theta \mathrm{I})^{2\nu+1}}{(2\nu+1)!},$$

which may, since I is idempotent, obviously be written in the form

$$e^{i\theta \mathrm{I}} = \mathrm{E} - \mathrm{I} + \left[\sum_{\nu=0}^{+\infty} \frac{(i\theta)^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{(i\theta)^{2\nu+1}}{(2\nu+1)!} \right] \mathrm{I} = \mathrm{E} - \mathrm{I} + \left[\sum_{\nu=0}^{+\infty} \frac{(i\theta)^{\nu}}{\nu!} \right] \mathrm{I},$$

from which (80) follows at once. In an analogous manner it is possible to obtain the identity

$$e^{i\theta \mathrm{Q}} = \mathrm{E} \cos\theta + i\mathrm{Q} \sin\theta,$$

valid for arbitrary real or complex θ and for an arbitrary square root Q of the unit matrix E (cf. p. 585). From the commutability condition (63b) we obtain (62′) from which if we apply (80) for $\theta = \pi$ and $\mathrm{I} = \mathrm{B}$ we obtain

$$e^{i\mathrm{H}} = (\mathrm{E} - 2\mathrm{B}) \cdot e^{-\mathrm{S}} = e^{-\mathrm{S}} \cdot (\mathrm{E} - 2\mathrm{B}),$$

which is, because of (63a), certainly a real matrix. q. e. d.

III. A Connection Between Reduced Skew-Symmetric and Reduced Hermitian Matrices.

For convenience we adopt the notation

$$\mathrm{M} \sim \mathrm{N} \tag{82}$$

for the orthogonal equivalence of two matrices, i. e. for the existence of an orthogonal matrix O for which $\mathrm{M} = \mathrm{ONO}^{-1}$. We introduce the definition

A real skew-symmetric matrix S *will be said to be reduced if the spectrum of the Hermitian matrix* $i\mathrm{S}$ *is contained in the closed interval* $[-\pi, +\pi]$ *and if the end-points* $-\pi$ *and* $+\pi$ *are not contained in the point spectrum of* $i\mathrm{S}$.

The property of a real skew-symmetric matrix to be reduced is an obvious orthogonal invariant.

We now show

LEMMA I. *Let* B *be a real symmetric idempotent matrix and let* S *be a real bounded skew-symmetric matrix commutable with* B, *i. e.*

$$BS = SB, \tag{83}$$

then the necessary and sufficient conditions in order that the Hermitian matrix $\pi B + iS$ *be reduced* (*cf. p.* 588) *are*

$$(E - B)S = S(E - B) = \| 0 \|, \tag{84a}$$

(84b) S *be reduced.*

Since B is a real symmetric idempotent matrix we have, according to a theorem of Hilbert,† in the notation (82)

$$B \sim \| \kappa_{pq} \| \text{ where } \kappa_{pq} = 0 \text{ for } p \neq q \text{ and } \kappa_{pp} = 0 \text{ or } \kappa_{pp} = +1. \tag{85}$$

We designate by α' and β' the number of times 0 and $+1$ respectively occur in the sequence $\{\kappa_{pp}\}$ so that necessarily $\alpha' \geqq 0$, $\beta' \geqq 0$ and $\alpha' + \beta' = +\infty$ and distinguish the following four possibilities for the matrix $\| \kappa_{pq} \|$ in (85):

$$\begin{aligned}
&\text{I:} && \alpha' = +\infty, && \beta' = 2m && \text{for } 0 \leqq m < +\infty,\\
(86)\quad &\text{II:} && \alpha' = +\infty, && \beta' = 2m+1 && \text{for } 0 \leqq m < +\infty,\\
&\text{III:} && \alpha' = m, && \beta' = +\infty && \text{for } 0 \leqq m < +\infty,\\
&\text{IV:} && \alpha' = +\infty, && \beta' = +\infty,
\end{aligned}$$

and corresponding to these four possibilities for the matrix $\| \kappa_{pq} \|$ in (85) we have the following four canonical forms for the matrix B:

$$\begin{aligned}
&\text{I:}\quad B \sim B^{(1)} = \| b^{(1)}{}_{pq} \|, \text{ where } b^{(1)}{}_{pq} = 0 \text{ for } p \neq q,\\
&\qquad \text{and } b^{(1)}{}_{pp} = 1 \text{ for } p = 1, 2, \cdots, 2m < +\infty,\\
&\qquad\qquad b^{(1)}{}_{pp} = 0 \text{ for } p > 2m,\\
&\text{II:}\quad B \sim B^{(2)} = \| b^{(2)}{}_{pq} \|, \text{ where } b^{(2)}{}_{pq} = 0 \text{ for } p \neq q,\\
(87)\qquad &\qquad \text{and } b^{(2)}{}_{pp} = 1 \text{ for } p = 1, 2, \cdots, 2m+1 < +\infty,\\
&\qquad\qquad b^{(2)}{}_{pp} = 0 \text{ for } p > 2m+1,\\
&\text{III:}\quad B \sim B^{(3)} = \| b^{(3)}{}_{pq} \|, \text{ where } b^{(3)}{}_{pq} = 0 \text{ for } p \neq q,\\
&\qquad \text{and } b^{(3)}{}_{pp} = 0 \text{ for } p = 1, 2, \cdots, m < +\infty,\\
&\qquad\qquad b^{(3)}{}_{pp} = 1 \text{ for } p > m,\\
&\text{IV:}\quad B \sim B^{(4)} = \| b^{(4)}{}_{pq} \|, \text{ where } b^{(4)}{}_{pq} = 0 \text{ for } p \neq q,\\
&\qquad \text{and } b^{(4)}{}_{pp} = b^{(4)}{}_{p+1\,p+1} = 1 \text{ for } p = 1, 5, 9, \cdots,\\
&\qquad\qquad b^{(4)}{}_{pp} = b^{(4)}{}_{p+1\,p+1} = 0 \text{ for } p = 3, 7, 11, \cdots.
\end{aligned}$$

† See "Hi", p. 145.

Since the commutability property (83) assumed for the matrices B and S in Lemma I is invariant under simultaneous transformations of B and S we must have, simultaneously with the matrix transformation of the matrix B of Lemma I into each of the above canonical forms, the reduction of the skew-symmetric matrix S of Lemma I to one of the following types:

$$\text{I:}\quad \mathrm{S}\sim\mathrm{S}^{(1)}=\left\|\begin{array}{c|c} \boldsymbol{S}_{2m} & \\ \hline & \mathrm{S}_{\infty-2m} \end{array}\right\|,$$

$$\boldsymbol{S}_{2m}=\bar{\boldsymbol{S}}_{2m}=-\boldsymbol{S}'_{2m},\quad \mathrm{S}_{\infty-2m}=\bar{\mathrm{S}}_{\infty-2m}=-\mathrm{S}'_{\infty-2m},$$

$$\text{II:}\quad \mathrm{S}\sim\mathrm{S}^{(2)}=\left\|\begin{array}{c|c} \boldsymbol{S}_{2m+1} & \\ \hline & \mathrm{S}_{\infty-(2m+1)} \end{array}\right\|,$$

$$\boldsymbol{S}_{2m+1}=\bar{\boldsymbol{S}}_{2m+1}=-\boldsymbol{S}'_{2m+1},\quad \mathrm{S}_{\infty-(2m+1)}=\bar{\mathrm{S}}_{\infty-(2m+1)}=-\mathrm{S}'_{\infty-(2m+1)},$$

(88)

$$\text{III:}\quad \mathrm{S}\sim\mathrm{S}^{(3)}=\left\|\begin{array}{c|c} \boldsymbol{S}_{m} & \\ \hline & \mathrm{S}_{\infty-m} \end{array}\right\|,$$

$$\boldsymbol{S}_{m}=\bar{\boldsymbol{S}}_{m}=-\boldsymbol{S}'_{m},\quad \mathrm{S}_{\infty-m}=\bar{\mathrm{S}}_{\infty-m}=-\mathrm{S}'_{\infty-m},$$

$$\text{IV:}\quad \mathrm{S}\sim\mathrm{S}^{(4)}=\left\|\begin{array}{cccccccc} \boldsymbol{S}_{11} & \boldsymbol{Z} & \boldsymbol{S}_{13} & \boldsymbol{Z} & . & . & . & \\ \boldsymbol{Z} & \boldsymbol{S}_{22} & \boldsymbol{Z} & \boldsymbol{S}_{23} & . & . & . & \\ \boldsymbol{S}_{31} & \boldsymbol{Z} & \boldsymbol{S}_{33} & \boldsymbol{Z} & . & . & . & \\ \boldsymbol{Z} & \boldsymbol{S}_{42} & \boldsymbol{Z} & \boldsymbol{S}_{44} & . & . & . & \\ . & . & . & . & & & & \\ . & . & . & . & & & & \\ . & . & . & . & & & & \end{array}\right\|,$$

$$\boldsymbol{S}_{pq}=\bar{\boldsymbol{S}}_{pq}=-\boldsymbol{S}'_{qp}=\left\|\begin{array}{cc} \alpha_{pq} & \beta_{pq} \\ \gamma_{pq} & \delta_{pq} \end{array}\right\|,\qquad \boldsymbol{Z}=\left\|\begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array}\right\|.$$

in which the capital letters represent matrices and the vacant spaces of the remaining two "quadrants" of I, II and III are occupied entirely by zeros. For example, the matrix $\mathrm{S}^{(1)}=\| s^{(1)}{}_{pq} \|$ of (88) is a bounded real skew-symmetric matrix for which

$$(89)\qquad s^{(1)}{}_{pq}=0,\quad \text{for}\quad p=1,2,\cdots,2m,\ 0\leqq 2m<q;$$
$$\text{and for}\quad 0\leqq 2m<p,\ q=1,2,\cdots,2m.$$

We now prove the necessity of the conditions (84a), (84b) in Lemma I. As mentioned in connection with Theorem I the property of a Hermitian matrix to be reduced is an orthogonal invariant. Consequently the Hermitian matrices

$$(90)\qquad \pi\mathrm{B}^{(j)}+i\mathrm{S}^{(j)};\qquad (j=1,2,3,4),$$

constructed from (87) and (88), are all reduced. In order to show the necessity of the conditions (84a), (84b) of Lemma I we consider separately each of the four possible canonical forms for the matrix B of Lemma I as given in (87).

1. *Proof of the necessity of* (84a) *for matrices* B *of Class I.* From I of (88) we are enabled to write

$$\mathbf{S}^{(1)} = \mathbf{S}_a^{(1)} + \mathbf{S}_b^{(1)}, \tag{91}$$

where, in the schematic notations of (88), (89), we have placed

$$\mathbf{S}_a^{(1)} = \left\| \begin{array}{c|c} \boldsymbol{S}_{2m} & \\ \hline & \end{array} \right\|, \qquad \mathbf{S}_b^{(1)} = \left\| \begin{array}{c|c} & \\ \hline & \mathbf{S}_{\infty-2m} \end{array} \right\|. \tag{92}$$

so that the matrix pair $(\mathbf{S}_a^{(1)}, \mathbf{S}_b^{(1)})$ is separated (cf. p. 582). Since the Hermitian matrix $\pi\mathrm{B}^{(1)} + i\mathbf{S}^{(1)}$ is reduced [cf. (90) and adjoining] the domain $\Delta(\pi\mathrm{B}^{(1)} + i\mathbf{S}^{(1)})$ lies in the closed interval $[0, 2\pi]$ (cf. p. 582), i. e. from (91)

$$\begin{aligned} 0 \leqq \Phi\,[(\pi\mathrm{B}^{(1)} + i\mathbf{S}^{(1)})\,; x] &= \pi\Phi\,[\mathrm{B}^{(1)}\,; x] \\ &+ \Phi\,[i\mathbf{S}_a^{(1)}\,; x] + \Phi\,[i\mathbf{S}_b^{(1)}\,; x] \leqq 2\pi \quad \text{for} \quad |\,x\,| = 1. \end{aligned} \tag{93}$$

Since (93) holds for an arbitrary unit vector x we may choose in (93)

$$x_1 = x_2 = \cdots = x_{2m} = 0, \qquad \sum_{\nu=2m+1}^{+\infty} |\,x_\nu\,|^2 = 1, \tag{94}$$

so that, from (92) and the definition of $\mathrm{B}^{(1)}$ in (87), we derive from (93)

$$0 \leqq \Phi\,[i\mathbf{S}_b^{(1)}\,; x] \leqq 2\pi, \tag{95}$$

first, for only those unit vectors x of (94), and then, because of the peculiar definition of $\mathbf{S}_b^{(1)}$ given in (92), for an arbitrary unit vector x. Since the domain of any bounded Hermitian matrix of the type $i\mathrm{S}$, where S is a real skew-symmetric matrix, is necessarily symmetrical with respect to the origin (cf. p. 582) one must perforce replace (95) by

$$\Phi\,[i\mathbf{S}_b^{(1)}\,; x] \equiv 0 \quad \text{for} \quad |\,x\,| = 1, \tag{96}$$

and therefore from the uniqueness property of compound forms (cf. p. 580)

$$\mathbf{S}_b^{(1)} = \|\,0\,\|,$$

or from (91)

$$\mathbf{S}^{(1)} = \mathbf{S}_a^{(1)}. \tag{97}$$

From (87), (97), and (92) it is clear that the matrix pair $(\mathrm{E} - \mathrm{B}^{(1)}, \mathbf{S}^{(1)})$

is separated and consequently *a fortiori* perpendicular. Since the perpendicularity of a matrix pair is invariant with respect to simultaneous matrix transformations of the two matrices in the pair (cf. p. 583) the matrix pair (E — B, S) of Lemma I is itself perpendicular and the necessity of (84a) is accordingly demonstrated, under the temporary restriction that B be of Class I.

2. *Proof of the necessity of* (84b) *for matrices* B *of Class I.* We choose instead of (94) those unit vectors x for which

$$\sum_{\nu=1}^{2m} |\, x_\nu \,|^2 = 1, \qquad x_{2m+1} = x_{2m+2} = \cdots = 0, \tag{98}$$

so that, from (92) and the definition of $B^{(1)}$ in (87) we derive from (93)

$$0 \leqq \pi + \Phi\,[iS_a{}^{(1)}; x] \leqq 2\pi, \quad \text{i. e.} \quad -\pi \leqq \Phi\,[iS_a{}^{(1)}; x] \leqq +\pi, \tag{99}$$

first, for only those unit vectors of (98), and then because of the peculiar definition of $S_a{}^{(1)}$ given in (92), for an arbitrary unit vector x. Since, as has been mentioned in the introduction (cf. p. 581), the spectrum of any bounded matrix is contained in the domain of the matrix it is clear from (99) that the spectrum of $iS_a{}^{(1)}$ is contained in the closed interval $[-\pi, +\pi]$. Furthermore the matrix $E - B^{(1)}$ is a diagonal matrix the elements of which are 0 and $+1$ and forms with $S^{(1)}$ the separated matrix pair $(E - B^{(1)}, S^{(1)})$ [cf. (87), (97), (92)] so that from p. 584 and the fact that the Hermitian matrix $\pi B^{(1)} + iS^{(1)}$ is reduced [cf. (90) and adjoining] the points $-\pi$ and $+\pi$ cannot be in the point spectrum of $iS^{(1)}$, i. e. $S^{(1)}$ is a reduced skew-symmetric matrix. Since the property of a skew-symmetric matrix to be reduced is an orthogonal invariant the skew-symmetric matrix S of Lemma I is itself reduced and the necessity of (84b) is, under the temporary restriction that B be of type I of (87), accordingly demonstrated.

3. *Proof of the necessity of* (84a) *and* (84b) *for the matrices* B *of Classes II, III, and IV.* That (84a) and (84b) are necessary if the matrix B in Lemma I is of type II of (87) follows at once on replacing $2m$ by $2m + 1$ throughout the preceding demonstrations. In order to treat type III of (87) we may write, analogously to (91) and (92), from (88)

$$S^{(3)} = S_a{}^{(3)} + S_b{}^{(3)}, \tag{100}$$

where

$$S_a{}^{(3)} = \left\| \begin{array}{c:c} \mathbf{S}_m & \\ \hdashline & \end{array} \right\|, \qquad S_b{}^{(3)} = \left\| \begin{array}{c:c} & \\ \hdashline & \mathbf{S}_{\infty-m} \end{array} \right\|. \tag{101}$$

so that the matrix pair $(S_a^{(3)}, S_b^{(3)})$ is separated. Since the matrix $\pi B^{(3)} + iS^{(3)}$ is reduced [cf. (90) and adjoining] we accordingly have (cf. p. 582) from (100)

$$(102)\quad 0 \leqq \Phi\,[(\pi B^{(3)} + iS^{(3)}); x] = \pi\Phi\,[B^{(3)}; x] + \Phi\,[iS_a^{(3)}; x] + \Phi\,[iS_b^{(3)}; x] \leqq 2\pi \text{ for } |x| = 1,$$

and hence

$$(103)\quad 0 \leqq \pi + \Phi[iS_b^{(3)}; x] \leqq 2\pi \text{ for } x_1 = x_2 = \cdots x_m = 0, \sum_{\nu=m+1}^{+\infty} |x_\nu|^2 = 1,$$

$$(104)\quad 0 \leqq \Phi[iS_a^{(3)}; x] \leqq 2\pi \text{ for } \sum_{\nu=1}^{m} |x_\nu|^2 = 1, \quad x_{m+1} = x_{m+2} \cdots = 0,$$

so that one may deduce [cf. (95), (99) and adjoining] from (103) and (104) that the matrix pair $(E - B^{(3)}, S^{(3)})$ is separated and that the skew-symmetric matrix $S^{(3)}$ is reduced and then, as a consequence, that the matrix pair $(E - B, S)$ of Lemma I is perpendicular and that the skew-symmetric matrix S is reduced. In order to treat the remaining case in which the matrix B of Lemma I is of the type IV of (87) we write

$$(105)\qquad S^{(4)} = S_a^{(4)} + S_b^{(4)},$$

where, in the notations employed in (88), we understand

$$(106)\quad S_a^{(4)} = \left\| \begin{matrix} S_{11} & Z & S_{13} & Z & \cdots \\ Z & Z & Z & Z & \cdots \\ S_{31} & Z & S_{33} & Z & \cdots \\ Z & Z & Z & Z & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|, \qquad S_b^{(4)} = \left\| \begin{matrix} Z & Z & Z & Z & \cdots \\ Z & S_{22} & Z & S_{24} & \cdots \\ Z & Z & Z & Z & \cdots \\ Z & S_{42} & Z & S_{44} & \cdots \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \end{matrix} \right\|,$$

so that the matrix pair $(S_a^{(4)}, S_b^{(4)})$ is separated. Since the Hermitian matrix $\pi B^{(4)} + iS^{(4)}$ is reduced [cf. (90) and adjoining] we accordingly have (cf. p. 582) from (105)

$$(107)\quad 0 \leqq \Phi\,[(\pi B^{(4)} + iS^{(4)}; x] = \pi\Phi\,[B^{(4)}; x] + \Phi\,[iS_a^{(4)}; x] + \Phi\,[iS_b^{(4)}; x] \leqq 2\pi \quad \text{for} \quad |x| = 1,$$

and since x is an arbitrary unit vector in (107) we may put

(108)

$$x_1 = x_2 = x_5 = x_6 = x_9 = x_{10} = \cdots = 0;$$
$$|x_3|^2 + |x_4|^2 + |x_7|^2 + |x_8|^2 + \cdots = 1,$$

so that from (106) and the definition of $B^{(4)}$ in (87) we may derive from (107)

$$0 \leqq \Phi\,[iS_b^{(4)}; x] \leqq 2\pi, \tag{109}$$

first, for only those unit vectors x in (108), and then, because of the peculiar definition of $S_b^{(4)}$ given in (106), for an arbitrary unit vector x. A treatment of inequality (109) analogous to that given for (95) then yields

$$S_b^{(4)} = \| 0 \|,$$

or from (105)

$$S^{(4)} = S_a^{(4)}. \tag{110}$$

From (87), (110), and (106) it is clear that the matrix pair $(E - B^{(4)}, S^{(4)})$ is separated and consequently perpendicular. From this point on the proof of the necessity of (84a) for those matrices B of Lemma I of type IV in (87) is identical with the proof given in the treatment of type I and is accordingly not repeated.

In order to show the necessity of (84b) if the matrix B in Lemma I is of type IV in (87) we choose, instead of (108), those unit vectors x for which

$$\begin{gathered} |x_1|^2 + |x_2|^2 + |x_5|^2 + |x_6|^2 + |x_9|^2 + |x_{10}|^2 = \cdots = 1, \\ x_3 = x_4 = x_7 = x_8 = \cdots = 0 \end{gathered} \tag{111}$$

so that now, from (106) and the definition of $B^{(4)}$ in (87), the inequality (107) is replaced by

$$0 \leqq \pi + \Phi\,[iS_a^{(4)}; x] \leqq 2\pi, \quad \text{i. e.} \quad -\pi \leqq \Phi\,[iS_a^{(4)}; x] \leqq +\pi, \tag{112}$$

first, only for those unit vectors x in (111), and then, because of the peculiar definition of $S_a^{(4)}$ given in (106), for an arbitrary unit vector x. By employing the same argument used in the discussion of (99) we learn that the spectrum of $iS_a^{(4)}$ is contained in the closed interval $[-\pi, +\pi]$ and that neither of the end-points $-\pi$ and $+\pi$ are contained in the point spectrum of $iS_a^{(4)}$ (cf. p. 610). Finally it follows, just as in the treatment on p. 610, that the matrix S in Lemma I is a reduced skew-symmetric matrix.

4. *Sufficiency of the conditions* (84a) *and* (84b) *in Lemma I.* We now turn to the proof of the sufficiency of (84a) and (84b) in Lemma I. Since B is now by hypothesis a real symmetric idempotent matrix and S a real skew-symmetric matrix commutable with B [the commutability condition (83) is a necessary but not a sufficient condition for the perpendicularity of the matrix pair $(E - B, S)$], we may by *simultaneous* orthogonal matrix transformations obtain the canonical forms in (87) and (88) for the matrices B and S, respectively. As has been mentioned before, the assumptions in

(84a) and (84b) for the matrices B and S are orthogonal invariants and since the property of a Hermitian matrix to be reduced is likewise an orthogonal invariant it will be sufficient to prove the sufficiency for the Hermitian matrices in (90). We consider separately the four possibilities in (87) and begin with the treatment of type I. From the perpendicularity of the matrix pair $(E - B^{(1)}, S^{(1)})$ there follows from (87) and (88) that $S_b^{(1)} = \| 0 \|$ in (92), i. e. from (91), the matrix equation (97). Since $S^{(1)}$ is a reduced skew-symmetric matrix it follows from (97), (92) and (87) that $\pi B^{(1)} + iS^{(1)}$ is a reduced Hermitian matrix (cf. p. 584); from which the sufficiency of (84a) and (84b) in Lemma I follows [if B is of type I in (87)]. The treatment of the types II and III of (87) obviously proceeds entirely along the same lines and accordingly is not repeated here. The treatment of type IV of (87), while proceeding in the same general manner, is however more complicated and consequently is given in some detail. From the perpendicularity of the matrix pair $(E - B^{(4)}, S^{(4)})$ there follows from (87) and (88) that $S_b^{(4)} = \| 0 \|$ in (106), i. e. from (105), the matrix equation (110). Since $S^{(4)}$ is a reduced skew-symmetric matrix it follows from (110), (106), and (87) that first, the spectrum of $\pi B^{(4)} + iS^{(4)}$ is contained in the closed interval $[0, 2\pi]$ and second, from p. 584, that the end-point 2π is not contained in the point spectrum of $\pi B^{(4)} + iS^{(4)}$; from which the sufficiency of (84a) in Lemma I for the matrices B of the type IV in (87) follows. q. e. d.

III. The Parametric Representation of the Manifold of the Orthogonal Matrices.

We first employ Theorem I and Lemma I for the demonstration of the following two theorems:

Theorem II. *Every orthogonal matrix* O *may be represented in the form*

$$O = e^{i(\pi B + iS)} = e^{i\pi B} \cdot e^{-S} = e^{-S} \cdot e^{i\pi B}, \tag{113}$$

in which B *is a real symmetric idempotent matrix and* S *a real bounded skew-symmetric matrix commutable with* B, *i. e.*

$$B = \bar{B} = B', \; B^2 = B; \quad S = \bar{S} = -S', \tag{114a}$$

$$BS = SB. \tag{114b}$$

Furthermore this representation is possible and unique if the matrix pair

$(E - B, S)$ *is perpendicular and if the skew-symmetric matrix* S *is reduced, i. e. if*

$$(E - B)S = S(E - B) = \| 0 \|, \tag{115a}$$

(115b) S *is reduced.*

THEOREM III. *Every orthogonal matrix* O *may be represented in the form*

$$O = 2C - e^{-S}, \tag{116}$$

in which C *is a real symmetric idempotent matrix and* S *a bounded real skew-symmetric matrix such that the matrix pair* (C, S) *is perpendicular, i. e.*

$$C = \bar{C} = C', \; C^2 = C; \quad S = \bar{S} = -S', \tag{117a}$$

$$CS = SC = \| 0 \|, \tag{117b}$$

and this representation is unique if the skew-symmetric matrix S *in* (116) *is reduced. The matrix* C *is in fact the matrix* E — B *of Theorem II.*

Theorem II is a trivial consequence of Theorem I and Lemma I in that for its demonstration it is only necessary to replace the condition in Theorem I that the Hermitian matrix $\pi B + iS$ be reduced by the equivalent conditions (84a) and (84b) of Lemma I. That the representation (113), (114a), (114b) always yields an orthogonal matrix, i. e. a real unitary matrix even without the conditions (115a), (115b) is trivial inasmuch as (cf. p. 606)

$$O = e^{i(\pi B + iS)} = e^{i\pi B} \cdot e^{-S} = (E - 2B)e^{-S} = \bar{O},$$

from which it is clear that the prescribing of the conditions (115a) and (115b) is responsible for the uniqueness of the representation (113).

Theorem III is now demonstrated as a consequence of Theorem II and the identity (19) on p. 588 valid for every perpendicular pair (M, N) of bounded matrices. In order to show this we write the representation (113) as follows:

$$\begin{aligned} O = \exp[i(\pi B + iS)] &= \exp[i(\pi E - \pi(E - B) + iS)] \\ &= -\exp i[-\pi(E - B) + iS]. \end{aligned} \tag{118}$$

If we now choose the matrices B and S to fulfill the conditions (114a), (114b), (115a), (115b) the representation (113), (118) provides a unique representation of any given orthogonal matrix O and which may, from (19), be written in the form

$$O = E - e^{-i\pi(E-B)} - e^{-S},$$

which in turn, from (80) [$\theta = -\pi$ and $I = E - B$], may be written as

$$O = 2(E - B) - e^{-S}, \tag{119}$$

inasmuch as for $B^2 = B$, as is well known,

$$(E - B)^2 = E^2 - 2B + B^2 = E - B.$$

The matrix $E - B$ in (119) may accordingly be identified as the matrix C in Theorem III. Finally it is clear that the representation (116), (117a), (117b) always yields an orthogonal matrix, even though S is not a reduced skew-symmetric matrix, for

$$\begin{aligned} OO' = O'O &= [2(E - B) - e^{-S}]\,[2(E - B) - e^{S}] \\ &= 4(E - B)^2 - 2(E - B)e^{S} - 2e^{-S}(E - B) + E \\ &= 4(E - B) - 2(E - B) - 2(E - B) + E = E. \end{aligned}$$

It is clear from the above considerations that the only effect on the representation (116) of prescribing that the skew-symmetric matrix S be reduced is to insure the uniqueness of the representation.

We now show

LEMMA II. *The formula* $Q = E - 2I$ *yields a one-to-one correspondence between the idempotent matrices* I *and the matrices* Q *which are square roots of* E. *This one-to-one correspondence has the obvious property that the matrices* I *and* Q *in it are simultaneously real and simultaneously symmetric.*

Let I be a given arbitrary (not necessarily Hermitian) idempotent matrix, i. e. $I^2 = I$. The matrix $Q = E - 2I$ is then uniquely determined and satisfies the matrix equation

$$Q^2 = (E - 2I)^2 = E^2 - 4I + 4I^2 = E,$$

so that Q is a square root of E.

Let Q be an arbitrary (not necessarily Hermitian) square root of E, i. e. $Q^2 = E$. The matrix $I = \frac{1}{2}(E - Q)$ is then uniquely determined and satisfies the matrix equation

$$I^2 = (1/4)(E - Q)^2 = (1/4)(E^2 - 2Q + Q^2) = (1/2)(E - Q) = I,$$

so that I is an idempotent matrix.

That the correspondence $Q = E - 2I$ is one to one is trivial.

As a consequence of Theorem I and Lemma II we show

THEOREM IV. *Every orthogonal matrix* O *may be represented in the form*

$$O = Qe^{-S} = e^{-S}Q, \tag{120}$$

in which Q *is a real symmetric square root of* E (*and therefore itself an orthogonal matrix inasmuch as* $QQ' = Q'Q = Q^2 = E$), S *a bounded real skew-symmetric matrix, commutable with* Q, *i. e.*

$$Q = \bar{Q} = Q', \; Q^2 = E; \quad S = \bar{S} = -S', \tag{121a}$$

$$QS = SQ, \tag{121b}$$

and this correspondence between the orthogonal matrices O *and the matrix pairs* (Q, S) *for which the Hermitian matrices* $(\pi/2)(E - Q) + iS$ *are reduced is one-to-one.*

The matrix B in the representation (62) is, from (63a), a real symmetric idempotent matrix. According to the above lemma there exists a uniquely determined real symmetric square root Q of E for which (62) may be written in the form

$$O = \exp i[(\pi/2)(E - Q) + iS] = \exp[i(\pi/2)(E - Q)] \cdot \exp[-S]. \tag{122}$$

The matrix $B = \frac{1}{2}(E - Q)$ is idempotent so that (122) may, from (80) [$\theta = \pi$ and $I = (1/2)(E - Q)$], be written in the desired form (120). Finally from Lemma II and Theorem I the correspondence (120), (122) between the orthogonal matrices O and the matrix pairs (Q, S) of (121a), (121b) is one-to-one provided that the Hermitian matrix $\pi B + iS$, i. e. $(\pi/2)(E - Q) + iS$ is reduced. q. e. d.

V. The Four Classes of Orthogonal Matrices.

According to Lemma II if Q is a real symmetric square root of E there exists a uniquely determined real symmetric idempotent matrix B for which

$$Q = E - 2B \tag{123}$$

holds. Corresponding to the four canonical forms for the matrices B, under orthogonal matrix transformations, given in (87) we accordingly obtain the following four canonical forms for the matrices Q which are square roots of E:

$$\begin{aligned}
&\text{I:} \quad Q \sim Q^{(1)} = \| q^{(1)}{}_{jk} \|, \; q^{(1)}{}_{jk} = 0 \; j \neq k, \; q^{(1)}{}_{jj} = -1 \\
&\qquad \text{for } j = 1, 2, \cdots, 2m < +\infty, \; q^{(1)}{}_{jj} = +1 \text{ for } j > 2m, \\
&\text{II:} \quad Q \sim Q^{(2)} = \| q^{(2)}{}_{jk} \|, \; q^{(2)}{}_{jk} = 0 \; j \neq k, \; q^{(2)}{}_{jj} = -1 \\
&\qquad \text{for } j = 1, 2, \cdots, (2m+1) < +\infty, \; q^{(2)}{}_{jj} = +1 \text{ for } j > 2m+1, \\
&\text{III:} \quad Q \sim Q^{(3)} = \| q^{(3)}{}_{jk} \|, \; q^{(3)}{}_{jk} = 0 \; j \neq k, \; q^{(3)}{}_{jj} = +1 \\
&\qquad \text{for } j = 1, 2, \cdots, m < +\infty, \; q^{(3)}{}_{jj} = -1 \text{ for } j > m,
\end{aligned} \tag{124}$$

$$\text{IV:}\quad Q \sim Q^{(4)} = \| q^{(4)}{}_{jk} \|,\ q^{(4)}{}_{jk} = 0\ j \neq k,\ q^{(4)}{}_{jj} = q_{j+1\,j+1} = -1$$
$$\text{for } j = 1, 5, 9, \cdots,\ q^{(4)}{}_{jj} = q^{(4)}{}_{j+1\,j+1} = +1 \text{ for } j = 3, 7, 11, \cdots,$$

On pages 607-612, we have shown that simultaneously with the transformation of the matrix B of Theorem I to a given canonical form (87) and hence simultaneously with the reduction of the matrix Q, defined by (123), to a canonical form in (124), the skew-symmetric matrix S of Theorem I is transformed into one of the four following forms:

$$\text{I:}\quad S \sim S^{(1)} = \left\| \begin{array}{c|c} \boldsymbol{S}_{2m} & \\ \hline & \end{array} \right\|,\quad \boldsymbol{S}_{2m} = \bar{\boldsymbol{S}}_{2m} = -\boldsymbol{S}'_{2m},$$

$$\text{II:}\quad S \sim S^{(2)} = \left\| \begin{array}{c|c} \boldsymbol{S}_{2m+1} & \\ \hline & \end{array} \right\|,\quad \boldsymbol{S}_{2m+1} = \bar{\boldsymbol{S}}_{2m+1} = -\boldsymbol{S}'_{2m+1},$$

$$\text{III:}\quad S \sim S^{(3)} = \left\| \begin{array}{c|c} & \\ \hline & S_{\infty-m} \end{array} \right\|,\quad S_{\infty-m} = \bar{S}_{\infty-m} = -S'_{\infty-m},$$

(125)

$$\text{IV:}\quad S \sim S^{(4)} = \left\| \begin{array}{cccc} \boldsymbol{S}_{11} & \boldsymbol{Z} & \boldsymbol{S}_{13} & \boldsymbol{Z} \ \cdots \\ \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} \ \cdots \\ \boldsymbol{S}_{31} & \boldsymbol{Z} & \boldsymbol{S}_{33} & \boldsymbol{Z} \ \cdots \\ \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} & \boldsymbol{Z} \ \cdots \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array} \right\|,$$

$$\boldsymbol{S}_{pq} = \bar{\boldsymbol{S}}_{pq} = -\boldsymbol{S}'_{qp} = \left\| \begin{array}{cc} \alpha_{pq} & \beta_{pq} \\ \gamma_{pq} & \delta_{pq} \end{array} \right\|,\quad \boldsymbol{Z} = \left\| \begin{array}{cc} 0 & 0 \\ 0 & 0 \end{array} \right\|.$$

[cf. for example for I of (125) formulas (91), (92), and (97)]. Since, according to Theorem IV, we may express any given orthogonal matrix O uniquely in the form (120) it is possible to effect a division of the orthogonal matrices into four classes, namely according as

(126) Class I: $O \sim Q^{(1)} e^{-S^{(1)}}$, Class II: $O \sim Q^{(2)} e^{-S^{(2)}}$,
Class III: $O \sim Q^{(3)} e^{-S^{(3)}}$, Class IV: $O \sim Q^{(4)} e^{-S^{(4)}}$,

in which the matrices $Q^{(i)}$, $S^{(i)}$ $(i = 1, 2, 3, 4)$ are defined in (124) and (125). As has been mentioned in connection with Theorem I this classification is possible because of the invariance of the spectra and the "covariance" property there pointed out with respect to orthogonal matrix transformations.

We now confine our attention to orthogonal matrices of Classes I and II in (126) and obtain a further simplification of the canonical forms in

(126) for these two classes. We shall understand by a canonical l-rowed skew-symmetric matrix $\boldsymbol{F}_l$ $(l < +\infty)$ the skew-symmetric matrix obtained by "inserting" in succession the real binary skew-symmetric matrices

$$(127)\quad \boldsymbol{S}_{pp} = \left\| \begin{matrix} 0 & s_p \\ -s_p & 0 \end{matrix} \right\|, \quad 1 \leqq p \leqq l/2 \ (l \text{ even}); \ 1 \leqq p \leqq (l-1)/2 \ (l \text{ odd}), \quad -\infty < s_p < +\infty$$

in the diagonal positions of the l-rowed square zero matrix. For even l every row and column of $\boldsymbol{F}_l$ contains one and only one s_l, the same being true for odd l with the exception of the first row and column which contain only zeros. As is well known every finite real skew-symmetric matrix $\boldsymbol{S}$ is orthogonally equivalent to a canonical skew-symmetric matrix so that the points in the spectrum of $\boldsymbol{S}$ are then necessarily

$$(128)\quad \pm is_1, \pm is_2, \cdots, \pm is_{l/2} \quad \text{or} \quad 0, \pm is_1, \pm is_2, \cdots, is_{(l-1)/2},$$

according as l is even or odd respectively. In an analogous manner we shall understand by an infinite canonical skew-symmetric matrix F_∞ the skew-symmetric matrix obtained by inserting infinitely many of the binary skew-symmetric matrices (127) in the diagonal positions of the infinite zero matrix. Those bounded skew-symmetric matrices, possessing no continuous spectrum, are analogous to finite real skew-symmetric matrices in that they are orthogonally equivalent † to an infinite canonical skew-symmetric matrix F_∞. With the use of canonical skew-symmetric matrices, as defined above, the canonical forms for the matrices of Classes I and II as set forth in (126) may be further simplified to

$$(129)\quad \text{Class I}: \mathrm{O} \sim \mathrm{Q}^{(1)} e^{-\mathrm{S}^{(1)}} \sim \mathrm{Q}^{(1)} e^{\mathrm{F}_{2m}}, \quad \text{Class II}: \mathrm{O} \sim \mathrm{Q}^{(2)} e^{-\mathrm{S}^{(2)}} \sim \mathrm{Q}^{(2)} e^{\mathrm{F}_{2m+1}},$$

inasmuch as, as is clear from I and II of (126) that

$$\mathrm{Q}^{(1)} e^{-\mathrm{S}^{(1)}} = \mathrm{Q}^{(1)} e^{-\mathrm{S}_{2m}}, \qquad \mathrm{Q}^{(2)} e^{-\mathrm{S}^{(2)}} = \mathrm{Q}^{(2)} e^{-\mathrm{S}_{2m+1}}.$$

By employing the identity ‡

† See "Hi", p. 163, where the above fact is demonstrated for completely continuous skew-symmetric matrices the proof permitting an immediate generalization to bounded skew-symmetric matrices, not possessing continuous spectra.

‡ The identity (130) may be derived on applying the Cayley-Hamilton theorem to the definition of exp M as given on p. 588. On placing

$$\boldsymbol{F}_\theta = \left\| \begin{matrix} 0 & \theta \\ -\theta & 0 \end{matrix} \right\|,$$

the Cayley-Hamilton theorem yields the recursion formulas

$$\exp \begin{Vmatrix} 0 & \theta \\ -\theta & 0 \end{Vmatrix} = \begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix} \tag{130}$$

and if l is odd also

$$\exp \begin{Vmatrix} 0 & 0 & 0 \\ 0 & 0 & \theta \\ 0 & -\theta & 0 \end{Vmatrix} = \begin{Vmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{Vmatrix},$$

in which θ denotes an arbitrary real or complex number, and the finite canonical l-rowed rotation matrices $\boldsymbol{O}_l$ defined by

$$\boldsymbol{O}_l = e^{\boldsymbol{F}_l}, \tag{131}$$

we are enabled to write (129) in the form

$$\text{Class I: } O \sim Q^{(1)}\boldsymbol{O}_{2m}, \qquad \text{Class II: } O \sim Q^{(2)}\boldsymbol{O}_{2m+1}, \tag{132}$$

and consequently replace the classification of the orthogonal matrices O as given in (126) by

$$\begin{aligned} \text{Class I:} \quad & O \sim Q^{(1)}e^{-S^{(1)}} \sim Q^{(1)}\boldsymbol{O}_{2m}, \\ \text{Class II:} \quad & O \sim Q^{(2)}e^{-S^{(2)}} \sim Q^{(2)}\boldsymbol{O}_{2m+1}, \\ \text{Class III:} \quad & O \sim Q^{(3)}e^{-S^{(3)}}, \\ \text{Class IV:} \quad & O \sim Q^{(4)}e^{-S^{(4)}}. \end{aligned} \tag{133}$$

For finite orthogonal matrices there exist no classes which would correspond to Classes III and IV of (133) whereas Classes I and II correspond to the finite orthogonal matrices of determinants $+1$ and -1 respectively since, denoting by $\det \boldsymbol{M}$ the determinant of an arbitrary finite matrix $\boldsymbol{M}$, we have from † (124)

$$\boldsymbol{F}_\theta^{2j} = (-1)^j\theta^{2j}\boldsymbol{E}, \quad \boldsymbol{F}_\theta^{2j+1} = (-1)^j\theta^{2j}\boldsymbol{F}_\theta; \qquad (j = 0, 1, 2, \cdots),$$

from which the matrix

$$e^{\boldsymbol{F}\theta} = \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{\nu}}{\nu!} = \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{2\nu}}{(2\nu)!} + \sum_{\nu=0}^{+\infty} \frac{\boldsymbol{F}_\theta^{2\nu+1}}{(2\nu+1)!},$$

takes the form

$$\begin{aligned} e^{\boldsymbol{F}\theta} &= \boldsymbol{E}\sum_{\nu=0}^{+\infty}(-1)^\nu \frac{\theta^{2\nu}}{(2\nu)!} + \frac{\boldsymbol{F}_\theta}{\theta}\sum_{\nu=0}^{+\infty}(-1)^\nu \frac{\theta^{2\nu+1}}{(2\nu+1)!} \\ &= \boldsymbol{E}\cos\theta + \frac{\sin\theta}{\theta}\boldsymbol{F}_\theta = \begin{Vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{Vmatrix}. \end{aligned}$$

Cf. the paper of Galanti referred to on p. 594.

† The first of the equations in (134) follows from the invariance of a determinant of a matrix with respect to matrix transformations, i. e.

$$\det e^{-\boldsymbol{S}} = \det \boldsymbol{O}e^{-\boldsymbol{S}}\boldsymbol{O}^{-1} = \det e^{-\boldsymbol{O}\boldsymbol{S}\boldsymbol{O}^{-1}},$$

and the easily verified fact that the finite canonical rotation matrices defined in (131) above are of determinant $+1$.

$$\begin{aligned} &\det e^{-\mathbf{S}} = (+1), \qquad \boldsymbol{S} = \bar{\boldsymbol{S}} = -\boldsymbol{S}', \\ &\det \boldsymbol{Q}^{(1)} = (-1)^{2m} = +1, \\ &\det \boldsymbol{Q}^{(2)} = (-1)^{2m+1} = -1. \end{aligned} \tag{134}$$

VI. Rotations and Reflections.

We now return to infinite orthogonal matrices and propose the definition: *We designate an orthogonal matrix* O *as a rotation matrix if and only if there exists a bounded real skew-symmetric matrix* G (*not necessarily unique*) *for which*

$$\mathrm{O} = e^{\mathrm{G}}, \tag{135}$$

holds. An orthogonal matrix not permitting this representation we designate as a reflection matrix.

In order to justify the introduction of this definition we show † the following lemma pertaining to finite orthogonal matrices:

Lemma III. *Every orthogonal matrix of determinant* $+1$ (*i. e. a finite rotation matrix*) *permits the representation* (135) *and no orthogonal matrix of determinant* -1 (*i. e. a finite reflection matrix*) *permits this representation.*

Let $\boldsymbol{O}$ be a given orthogonal matrix of determinant $+1$, i. e. the matrix $\boldsymbol{O}$ is of Class I in (133) so that we have

$$\boldsymbol{O} \sim \boldsymbol{Q}^{(1)} e^{\boldsymbol{F}_{2m}}, \tag{136}$$

in which the matrices $\boldsymbol{Q}^{(1)}$ and $\boldsymbol{F}_{2m}$ are defined on p. 616 and p. 618, respectively. If we now introduce a canonical skew-symmetric matrix $\boldsymbol{F}^{(1)}{}_{2m}$ defined by placing

† That every orthogonal matrix of determinant $+1$ permits the representation (135) is not, as appears at first sight, a trivial consequence of the normal forms for orthogonal matrices under orthogonal matrix transformations and the above identity (130) is clear from the example

$$\left\| \begin{matrix} \frac{\pi^2 - 2b^2}{\pi^2} & -\frac{2b(\pi^2 - b^2)^{\frac{1}{2}}}{\pi^2} & 0 \\ -\frac{2b(\pi^2 - b^2)^{\frac{1}{2}}}{\pi^2} & -\frac{\pi^2 - 2b^2}{\pi^2} & 0 \\ 0 & 0 & -1 \end{matrix} \right\| = \exp \left\| \begin{matrix} 0 & 0 & b \\ 0 & 0 & (\pi^2 - b^2)^{\frac{1}{2}} \\ -b & -(\pi^2 - b^2)^{\frac{1}{2}} & 0 \end{matrix} \right\|$$

of an orthogonal matrix of determinant $+1$ which is already in canonical form and whose exponential representation (135) is not at all obvious from the identity (130).

$$s_i = 1 \qquad (i = 1, 2, \cdots),$$

in (128) we are enabled to write the matrix $\boldsymbol{Q}^{(1)}$ of (136), (124) in the exponential form

$$\boldsymbol{Q}^{(1)} = e^{\pi \boldsymbol{F}^{(1)}{}_{2m}}, \tag{137}$$

inasmuch as, as follows from (130) for $\theta = \pi$, we have

$$\exp \pi \left\| \begin{matrix} 0 & 1 \\ -1 & 0 \end{matrix} \right\| = \left\| \begin{matrix} -1 & 0 \\ 0 & -1 \end{matrix} \right\|.$$

Since the matrices $\boldsymbol{F}_{2m}$ and $\boldsymbol{F}^{(1)}{}_{2m}$ are obviously commutable we may employ the distributive property of the exponential representation with respect to commutable matrices and write (136) with the help of (137) in the form

$$\boldsymbol{O} \sim e^{\pi \boldsymbol{F}^{(1)}{}_{2m}} \cdot e^{\boldsymbol{F}_{2m}} = \exp\left[\pi \boldsymbol{F}^{(1)}{}_{2m} + \boldsymbol{F}_{2m}\right],$$

and hence, since the property of being a rotation matrix in the sense of our definition is for finite as well as for infinite matrices obviously an orthogonal invariant,† the demonstration of the first part of Lemma III is completed.

The demonstration of the second part of Lemma III, namely that every finite matrix permitting the representation (133) is necessarily of determinant $+1$ is contained in (134) (cf. footnote thereto on page 619).

As an application of the exponential definition of rotation and reflection matrices as given in (135) we show

THEOREM V. *Every orthogonal matrix of Class I is, in the sense of the definition* (135), *a rotation matrix and every orthogonal matrix of Class II is, in the sense of definition* (135), *a reflection matrix.*

The demonstration of the first part of this theorem is seen to be identical, from the formula (136) on, of the first part of Lemma III and consequently is not repeated. As has been mentioned above, the property of an orthogonal matrix to be a rotation or a reflection matrix, in the sense of the definition on p. 620, is an orthogonal invariant. Consequently in order to demonstrate that the matrices of Class II in (133) are all reflection matrices it will be sufficient to demonstrate that the matrices $\mathrm{Q}^{(2)}\boldsymbol{O}_{2m+1}$ of II in (133) are all reflection matrices. The second part of Theorem V may not, as appears at

† This invariance property is trivial inasmuch as follows from the definition of e^{G} on p. 588 that if O is an arbitrary orthogonal matrix then $\mathrm{O}e^{\mathrm{G}}\mathrm{O}^{-1} = e^{\mathrm{OGO}^{-1}}$ in which OGO^{-1} is again a bounded real skew-symmetric matrix.

first sight, be reduced to the fact that no finite orthogonal matrix of determinant -1 permits the representation (135) [cf. (130)], since it would be possible that $Q^{(2)}e^{F_{2m+1}} = Q^{(2)}O_{2m+1} = e^{G}$ in which G is a bounded real skew-symmetric matrix containing infinitely many elements different from zero. In order to give a correct demonstration we first show

LEMMA IV. *The spectrum of an arbitrary real bounded Hermitian matrix* H (*not necessarily reduced*) *is mapped upon the spectrum of the matrix* e^{iH} *by the transformation* $w = e^{iz}$, *i. e. if* θ (θ *real*) *is a point in the spectrum of H then* $e^{i\theta}$ *is a point in the spectrum of* e^{iH}.

The proof of this lemma is trivial for reduced Hermitian matrices as it follows directly from the connection between the spectral matrices of H and e^{iH} as given in (59). We shall however require the result of this lemma for an arbitrary Hermitian matrix H and accordingly give a direct proof valid even if the matrix is not reduced. We adopt the notations for an arbitrary Hermitian matrix M

$$(138)\qquad \begin{array}{ll} \boldsymbol{l}[M] & \text{greatest lower bound of } \Phi[M;x] \text{ for } |x| = 1, \\ \boldsymbol{u}[M] & \text{least upper bound of } |\Phi[M;x]| \text{ for } |x| = 1. \end{array}$$

We shall require the well known inequalities †

$$(139)\qquad |\Phi[MN;x]| \leqq (\Phi[MM^*;x]\,\Phi[N^*N;x])^{\frac{1}{2}} \quad \text{for} \quad |x| = 1,$$
$$(140)\qquad \boldsymbol{u}[M,N] \leqq \boldsymbol{u}[M]\,\boldsymbol{u}[N],$$
$$(141)\qquad \boldsymbol{l}[MM^*] \geqq 0 \quad \boldsymbol{l}[M^*M] \geqq 0,$$

valid for any two bounded matrices M and N.

First we derive the relations

$$(142)\qquad \begin{aligned} &\Phi[(\theta E - H)(\theta E - H)^*;x] \\ &\quad = \Phi[(\theta E - H)^*(\theta E - H);x] = \Phi[(\theta E - H)^2;x] \geqq 0, \end{aligned}$$

$$(143)\qquad \begin{aligned} &\Phi[(e^{i\theta}E - e^{iH})(e^{i\theta}E - e^{iH})^*;x] \\ &\quad = \Phi[(e^{i\theta}E - e^{iH})^*(e^{i\theta}E - e^{iH});x] = -2\sum_{\nu=1}^{+\infty}(-1)^{\nu}\frac{\Phi[(\theta E - H)^{2\nu};x]}{(2\nu)!} \end{aligned}$$

$$(144)\qquad 0 \leqq \Phi[(\theta E - H)^{2k};x] \leqq (\Phi[(\theta E - H)^2;x]\,\Phi[(\theta E - H)^{4k-2};x])^{\frac{1}{2}}; \quad (k = 1, 2, \cdots),$$

valid for real θ, an arbitrary unit vector x, and an arbitrary bounded Hermitian matrix H. The proof of (142) is, from (141), trivial. For the demonstration of (143) we observe that from

† Cf. for instance "S", p. 137, p. 130, p. 132, respectively.

$$(145)\quad \Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})^*;x] = \Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})(e^{-i\theta}\mathrm{E}-e^{-i\mathrm{H}});x] = \Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})^*(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}});x],$$

we obtain

$$(146)\quad \Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})^*;x] = 2-\Phi[(e^{i(\theta\mathrm{E}-\mathrm{H})}+e^{-i(\theta\mathrm{E}-\mathrm{H})});x],$$

from which (143) obviously follows. The inequalities (144) are derived from that of (139) by placing therein $\mathrm{M}=(\theta\mathrm{E}-\mathrm{H})$, $\mathrm{N}=(\theta\mathrm{E}-\mathrm{H})^{2k-1}$ and $k=2,3\cdots$ in succession and at the same time employing (142) and

$$0\leqq\Phi[(\theta\mathrm{E}-\mathrm{H})^2(\theta\mathrm{E}-\mathrm{H})^{2*};x]=\Phi[(\theta\mathrm{E}-\mathrm{H})^4;x].$$

If H is a diagonal matrix the lemma is trivial and we exclude this case in what follows. Let θ be a point in the spectrum of H, i. e. assume [cf. (581)]

$$(147)\quad \boldsymbol{l}[(\theta\mathrm{E}-\mathrm{H})(\theta\mathrm{E}-\mathrm{H})^*] = \boldsymbol{l}[(\theta\mathrm{E}-\mathrm{H})^*(\theta\mathrm{E}-\mathrm{H})]=\boldsymbol{l}[(\theta\mathrm{E}-\mathrm{H})^2]=0,$$

so that there exists a sequence of unit vectors $\{x_n\}$ for which

$$(148)\quad \lim_{n\to+\infty}\Phi[(\theta\mathrm{E}-\mathrm{H})^2;x_n]=0.$$

From (142), (143), and (144) we have

$$(149)\quad 0\leqq\Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})^*;x]\leqq 2\sum_{\nu=1}^{+\infty}\frac{\Phi[(\theta\mathrm{E}-\mathrm{H})^{2\nu},x]}{(2\nu)!}$$
$$\leqq 2(\Phi[(\theta\mathrm{E}-\mathrm{H})^2;x])^{\frac{1}{2}}\sum_{\nu=1}^{+\infty}\frac{(\Phi[(\theta\mathrm{E}-\mathrm{H})^{4\nu-2};x])^{\frac{1}{2}}}{(2\nu)!}.$$

Since $\theta\mathrm{E}-\mathrm{H}\neq\|0\|$ is bounded we have

$$(150)\quad 0<\boldsymbol{u}[\theta\mathrm{E}-\mathrm{H}]=\boldsymbol{u},$$

where we write, for the sake of brevity, $\boldsymbol{u}[\theta\mathrm{E}-\mathrm{H}]=\boldsymbol{u}$ and hence have from (138) and (140)

$$(\Phi[(\theta\mathrm{E}-\mathrm{H})^{4\nu-2};x])^{\frac{1}{2}}\leqq(\boldsymbol{u}[(\theta\mathrm{E}-\mathrm{H})^{4\nu-2}])^{\frac{1}{2}}\leqq\boldsymbol{u}^{2\nu-1};\quad(\nu=1,2,\cdots),$$

so that the inequality (149) may be replaced by

$$(151)\quad 0\leqq\Phi[(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})(e^{i\theta}\mathrm{E}-e^{i\mathrm{H}})^*;x]$$
$$\leqq\frac{2}{\boldsymbol{u}}(\Phi[(\theta\mathrm{E}-\mathrm{H})^2;x])^{\frac{1}{2}}\sum_{\nu=1}^{+\infty}\frac{\boldsymbol{u}^{2\nu}}{(2\nu)!}<\frac{2e^{\boldsymbol{u}}}{\boldsymbol{u}}(\Phi[(\theta\mathrm{E}-\mathrm{H})^2;x])^{\frac{1}{2}}.$$

If one now applies the inequality (151) to the sequence of vectors $\{x_n\}$ for which (148) holds there follows from (151) the fulfillment of (3) for

$$\lambda = e^{i\theta}, \qquad \mathrm{M} = e^{i\mathrm{H}},$$

i. e. $e^{i\theta}$ is a point in the spectrum of $e^{i\mathrm{H}}$.

If S is a real bounded skew-symmetric matrix the matrix $i\mathrm{S}$ is a bounded Hermitian matrix so that we have the

COROLLARY. *The spectrum of an arbitrary real bounded skew-symmetric matrix* S *is mapped upon the spectrum of the matrix* e^{S} *by the transformation* $w = e^z$, *i. e. if* $i\theta$ *is a point in the spectrum of* S *then* $e^{i\theta}$ *is a point in the spectrum of* e^{S}.

We now demonstrate the second part of Theorem V and first calculate the matrix $\mathrm{Q}^{(2)}\boldsymbol{O}_{2m+1}$ of (133) obtaining from (124), (132), (127) and (131)

(152) $\mathrm{Q}^{(2)}\boldsymbol{O}_{2m+1}$

$$= \left\| \begin{array}{cccccccccc} -1 & & & & & & & & & \\ & -\cos s_1 & -\sin s_1 & & & & & & & \\ & \sin s_1 & -\cos s_1 & & & & & & & \\ & & & -\cos s_2 & -\sin s_2 & & & & & \\ & & & \sin s_2 & -\cos s_2 & & & & & \\ & & & & & \cdot & & & & \\ & & & & & & \cdot & & & \\ & & & & & & & -\cos s_m & -\sin s_m & \\ & & & & & & & \sin s_m & -\cos s_m & \\ & & & & & & & & & 1 \\ & & & & & & & & & & 1 \\ & & & & & & & & & & & \cdot \\ & & & & & & & & & & & & \cdot \end{array} \right\|,$$

from which it is clear that the matrix (152) possesses only a point spectrum namely the points

(153) $$-1,\ e^{\pm s_1},\ e^{\pm i s_2},\ \cdots,\ e^{\pm i s_m},\ 1,\ 1,\ \cdots$$

so that -1 necessarily occurs an odd number of times in the spectrum of (152). If now the matrix (152) were to permit the exponential representation e^{G} of (135) the matrix G therein could, from the above corollary, possess no continuous spectra and consequently (cf. p. 618) is orthogonally equivalent to an infinite canonical skew-symmetric matrix F_∞. Since the spectrum of any matrix is invariant under matrix transformation it is clear that either the points

(154) $$(2k+1)\pi i, \qquad -(2k+1)\pi i; \qquad (k = 0, 1, \cdots),$$

belong simultaneously to the spectrum of G or neither belongs to the spectrum

of G. Therefore if -1 occurs in the spectrum of e^G it follows from (154) and the above corollary that it must occur an even or an infinite number of times which is a contradiction to (153). q. e. d.

APPENDIX.

The Problem of moments for a finite interval and the Theory of Bounded Hermitian Matrices.

Several presentations have been given for the spectral representation of a bounded Hermitian matrix by means of Stieltjes integrals as introduced by Hilbert. The various modes of investigation require in common with the method of Hilbert, as far as the author knows from the literature,† the use of the characteristic constants of the truncated matrices of the given bounded Hermitian matrix. This procedure has certain methodical inconveniences. First, the spectra of the truncated matrices of a bounded Hermitian matrix do not necessarily determine the spectrum of the infinite matrix itself. An example of such a bounded Hermitian matrix has been pointed out by Toeplitz ‡ in which he considers the bounded Hermitian matrix constructed from the infinite zero matrix by inserting the real binary symmetric matrix

$$\begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}$$

in the diagonal positions of the infinite zero matrix, i. e. the matrix of the Hermitian form

$$x_1\bar{x}_2 + x_2\bar{x}_1 + x_3\bar{x}_4 + x_4\bar{x}_3 + \cdots,$$

and which contains only the points -1 and $+1$ in its spectrum whereas not only -1 and $+1$ but also 0 belong to the spectrum of infinitely many truncated matrices. Second, the method of the truncated matrices is restricted to the treatment of Hermitian matrices alone. For instance the truncated matrices of a unitary matrix are not necessarily also unitary matrices although it is possible to obtain by a direct treatment of the trigonometrical momentum problem belonging to the infinite matrix itself a spectral

† I. See "Hi", pp. 131-137; II. Hellinger Toeplitz, "Integralgleichungen und Gleichungen mit unendlichvielen Unbekannten," *Encyklopädie der Mathematische Wissenschaften* 2^{III_2}, pp. 1575-1584. III. F. Riesz, "Über quadratische Formen von unendlichvielen Veränderlichen," *Göttingen Nachrichten* (1910), pp. 190-195. Cf. also by the same author "R", pp. 128-139. We refer to the article of Hellinger Toeplitz as "HT" and this paper of Riesz as "U".

‡ See "T", p. 107.

representation which is entirely analogous to that discovered by Hilbert † for bounded Hermitian matrices. It is shown in the present note that such a direct treatment is also possible for bounded Hermitian matrices and yields a very simple demonstration of the fundamental theorem of Hilbert. It is necessary to start, not with the problem of trigonometrical moments but with the ordinary problem of moments for a finite interval as developed by Hausdorff.‡ The method possesses, in contrast to the methods of F. Riesz, not only the advantage of avoiding the use § of the truncated matrices at all but also enables us to start directly with a monotone Stieltjes kernel (Belegungsfunktion) whereas F. Riesz deduces ¶ the result from his general theorem ‖ on the solvability of a system of Stieltjes integral equations with a function of bounded variation and then has to prove that the solution is actually monotone. It is of course a matter of fact that the demonstration of the theorem of Hausdorff is based on the theorems of Helly applied by Carleman in connection with the method of the truncated matrices as developed by Hilbert. In addition to the theorem of Hausdorff we need the fact that every bounded positive definite matrix possesses a bounded "square root". If one is already in possession of the spectral representation of a bounded Hermitian matrix (the deduction of which is the purpose of this note) the uniqueness,†† as well as the existence of a bounded positive definite "square root" of a bounded positive definite matrix, is assured. We need, however, only the existence of a "square root" of a bounded positive definite matrix and this is available, according to a remark of eHllinger-Toeplitz,‡‡ without the use of a spectral representation.

For the sake of completeness we first discuss in some detail the remark of Hellinger-Toeplitz concerning the construction of a bounded positive definite square root of a bounded positive definite matrix. The basis of this remark is the ordinary Taylor development

---

† See "Hi", p. 138.

‡ See "Ha".

§ See "R", pp. 129-139, where Riesz gives a direct proof employing an approximation by matrix polynomials which however also depends upon the use of the truncated matrices.

¶ See "U".

‖ F. Riesz, "Sur les opérations fonctionelles linéaires," *Comptes Rendus*, 29 novembre (1909), and "Sur certains systèmes d'equations fonctionelles et l'approximation des fonctions continues," *Comptes Rendus*, 14 mars (1910).

†† See "W", p. 267. For finite matrices a direct demonstration of this existence and uniqueness property is given by L. Autonne in "Sur l'Hermitien," *Rendiconti del Circolo Matematico di Palermo*, Vol. 16 (1902), pp. 120-121.

‡‡ See "HT", p. 1567, footnote (522 a).

$$f(z) = (1-z)^{\frac{1}{2}} = \sum_{\nu=0}^{+\infty} a_\nu z^\nu, \qquad |z| < 1, \tag{155}$$

in which

$$a_0 a_0 = 1, \quad a_0 a_1 + a_1 a_0 = -1, \quad \sum_{\mu=0}^{n} a_\mu a_{n-\mu} = 0; \qquad (n = 2, 3, \cdots). \tag{156}$$

We denote by

$$\Psi[\mathrm{A}; x, y] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p y_q, \tag{157}$$

the bilinear form of an arbitrary bounded matrix A and by

$$\Phi[\mathrm{A}; x] = \Psi[\mathrm{A}; x, \bar{x}] = \sum_{p=1}^{+\infty} \sum_{q=1}^{+\infty} a_{pq} x_p \bar{x}_q, \tag{158}$$

the compound form (Kopplungsform) of A. The upper limit of $|\Psi[\mathrm{A}; x, y]|$ for those vectors x and y for which

$$|x| = (\sum_{\nu=1}^{+\infty} |x_\nu|^2)^{\frac{1}{2}} = 1, \qquad |y| = (\sum_{\nu=1}^{+\infty} |y_\nu|^2)^{\frac{1}{2}} = 1, \tag{159}$$

will be denoted by $\boldsymbol{R}(\mathrm{A})$, the upper and lower limits of $|\Phi[\mathrm{A}; x]|$ for those vectors x of (159) being denoted by $\boldsymbol{M}(\mathrm{A})$ and $\boldsymbol{m}(\mathrm{A})$ respectively. It is known that if A_1 and A_2 are two arbitrary bounded matrices then according to the inequality of Schwarz †

$$\boldsymbol{R}(\mathrm{A}_1\mathrm{A}_2) \leqq \boldsymbol{R}(\mathrm{A}_1)\boldsymbol{R}(\mathrm{A}_2). \tag{160}$$

Furthermore it is clear that the series

$$\sum_{\nu=0}^{+\infty} \mathrm{A}_\nu \tag{161}$$

converges to a bounded matrix A if

$$\sum_{\nu=0}^{+\infty} \boldsymbol{R}(\mathrm{A}_\nu) < +\infty. \tag{162}$$

For an arbitrary bounded Hermitian matrix H it is known that the compound form $\Phi[\mathrm{H}; x]$ takes only real values and that

$$\boldsymbol{R}(\mathrm{H}) = \boldsymbol{M}(\mathrm{H}). \tag{163}$$

A bounded Hermitian matrix for which the upper and lower limits of its compound form on the complex Hilbert sphere $|x| = 1$ of (159) are con-

† Cf. for instance "S", p. 130.

tained in the open interval $(0, 1)$ will be denoted by K. We accordingly have for every Hermitian matrix K

$$0 < \boldsymbol{m}(\mathrm{K}) \leqq \boldsymbol{M}(\mathrm{K}) < 1, \tag{164}$$

in which the equality sign holds for and only for those matrices $\gamma \mathrm{E}$ in which E is the unit matrix and $0 < \gamma < 1$. The matrix $\mathrm{E} - \mathrm{K}$ is likewise a bounded Hermitian matrix for which we have from (164)

$$0 < \boldsymbol{m}(\mathrm{E} - \mathrm{K}) \leqq \boldsymbol{M}(\mathrm{E} - \mathrm{K}) < 1. \tag{165}$$

The matrix

$$\mathrm{A} = \sum_{\nu=0}^{+\infty} a_\nu K^\nu, \tag{166}$$

in which the a_ν are the coefficients in the Taylor development (155) and consequently fulfill the bilinear relations (156), is a bounded matrix since, from (164) and the definition of the a_ν in (155), (156) we have from (160) and (163)

$$\sum_{\nu=0}^{+\infty} \boldsymbol{R}(a_\nu K^\nu) = \sum_{\nu=0}^{+\infty} | a_\nu | \, \boldsymbol{R}(K^\nu) \leqq \sum_{\nu=0}^{+\infty} | a_\nu | \, \boldsymbol{R}(K)^\nu = \sum_{\nu=0}^{+\infty} | a_\nu | \, \boldsymbol{M}(K)^\nu < + \infty,$$

and therefore

$$\boldsymbol{R}(\mathrm{A}) \leqq \sum_{\nu=0}^{+\infty} \boldsymbol{R}(a_\nu \mathrm{K}^\nu) < + \infty.$$

One then verifies easily from (166) and the definition of the a_ν in (155), (156) that

$$\mathrm{A}^2 = \mathrm{E} - \mathrm{K}, \tag{167}$$

and, in an analogous manner from (165), that the matrix

$$\mathrm{B} = \sum_{\nu=0}^{+\infty} a_\nu (\mathrm{E} - \mathrm{K})^\nu, \tag{168}$$

is bounded and fulfills the matrix equation

$$\mathrm{B}^2 = \mathrm{K}. \tag{169}$$

The matrices A and B defined in (166) and (168) may then be said to be "square roots" of the bounded positive definite matrices $\mathrm{E} - \mathrm{K}$ and K respectively. The matrices A and B in (166) and (168) are in addition positive definite but inasmuch as we shall not need this fact in the following we give no demonstration. The matrices A and B are obviously Hermitian and commutable, i. e.

$$\mathrm{AB} = \mathrm{BA} = \mathrm{A}^*\mathrm{B}^* = \mathrm{B}^*\mathrm{A}^* \quad \text{where} \quad \mathrm{A}^* = \bar{\mathrm{A}}', \ \mathrm{B}^* = \bar{\mathrm{B}}'. \tag{170}$$

We now employ the results of the previous paragraph in the demonstration † of the

LEMMA. *If* K *is any bounded Hermitian matrix whose compound form* $\Phi[K;x]$ *is bounded by the inequalities* (164), *i. e.*

$$0 < \theta_1 < \Phi[K;x] < \theta_2 < 1 \quad \text{for} \quad |x| = 1, \tag{171}$$

then the matrix polynomials

$$\pi_{n,m}(K) \quad \text{where} \quad \pi_{n,m}(x) = x^m(1-x)^{n-m}; \quad (m = 0, 1, \cdots, n); \ (n = 0, 1, \cdots), \tag{172}$$

are not-negative definite matrices, i. e.

$$0 \leqq \Phi[\pi_{m,n}(K);x] \quad \text{for} \quad |x| = 1; \quad (m = 0, 1, \cdots, n); \ (n = 0, 1, \cdots). \tag{173}$$

In order to demonstrate (173) we employ the "square root" representation (167), (169) for the matrices $E - K$ and K respectively in the matrix polynomials (172) to obtain with the aid of the multiplication theorems (Faltungsätze) of Hilbert and the commutability property (170)

$$\pi_{m,n}(K) = (B^2)^m(A^2)^{n-m} = (B^mA^{n-m})(B^mA^{n-m}) = (B^mA^{n-m})(B^mA^{n-m})^*; \quad (m = 0, 1, \cdots, n; \ n = 0, 1, \cdots),$$

from which (173) follows at once inasmuch as the Hermitian matrix $(B^mA^{n-m})(B^mA^{n-m})^*$ is known to be non-negative definite.

We now consider the problem of moments ‡ for a finite interval, viz.

$$c_l = \int_0^1 \mu^l d\rho(\mu); \qquad (l = 0, 1, \cdots), \tag{174}$$

---

† The demonstration of this lemma may also be seen as a consequence of a theorem given by F. Riesz in "R", p. 129, which states that for a Hermitian matrix H whose compound form satisfies the inequalities

$$m \leqq \Phi[H;x] \leqq M \quad \text{for} \quad |x| = 1,$$

the compound form of the matrix polynomial

$$f(H) \quad \text{where} \quad f(x) = \sum_{\nu=0}^{n} c_\nu x^\nu,$$

constructed from the matrix H, satisfies the inequalities

$$\lambda \leqq \Phi[f(H);x] \leqq \Lambda \quad \text{for} \quad |x| = 1,$$

where

$$\lambda \leqq f(\mu) \leqq \Lambda \quad \text{for} \quad m \leqq \mu \leqq M.$$

In the demonstration for this theorem, given by Riesz, use is made, however, of the truncated matrices of the matrix $f(H)$ whereas in the direct demonstration of the lemma given above, we avoid this procedure.

‡ See "Ha", pp. 222-232.

in which the c_l are preassigned real numbers and there is desired a real monotone non-decreasing function $\rho(\mu)$ of the real argument μ which satisfies the system (174) of Stieltjes integral equations. According to Hausdorff ‡ this problem possesses a solution then and only then if the sequence $\{c_l\}$ is completely monotone, i. e. if and only if the following inequalities hold:

$$(175) \qquad \Delta^{n-m}c_m \geqq 0; \qquad (m=0,1,\cdots,n);\ (n=0,1,\cdots),$$

where

$$\Delta c_m = c_m - c_{m+1}, \qquad \Delta^j = \Delta^{j-1}\Delta; \qquad (j=2,3,\cdots),$$

so that

$$\Delta^{n-m}c_m = c_m - \binom{n-m}{1}c_{m+1} + \cdots + (-1)^{n-m}c_n; \qquad (m=0,1,\cdots,n);\ (n=0,1,\cdots).$$

As is well known one may then insure the uniqueness of the solution $\rho(\mu)$ of (174) at all points of its interval of definition, the closed interval $[0,1]$, by prescribing, for instance, the normalizing conditions

$$(176) \qquad 0=\rho(0) \quad \text{and} \quad \rho(\mu-0)=\rho(\mu), \qquad 0<\mu<1,$$

for the solution $\rho(\mu)$ of (174), obtaining as a consequence of the assumption $\rho(0)=0$ from (174) for $l=0$

$$(177) \qquad \rho(1)-\rho(0)=\rho(1)=c_0.$$

For our purposes we consider the problem of moments arising from (174) if one places

$$(178) \qquad c_l=\Phi[\mathrm{K}^l;x], \quad |x|=1, \quad \mathrm{K}^0=\mathrm{E}; \qquad (l=0,1,\cdots),$$

in which x is an arbitrary but fixed unit vector and K a Hermitian matrix whose compound form $\Phi[\mathrm{K};x]$ is bounded by the inequalities (164), i. e. (171). We seek a monotone non-decreasing function $\rho(\mu;x)$ of the real argument μ and the components $x_1, x_2, \cdots$ of the above given unit vector x which satisfies the system (174), (178) of Stieltjes integral equations, i. e. is a solution of the problem of moments

$$(179) \qquad \Phi[\mathrm{K}^l;x]=\int_0^1 \mu^l\, d\rho(\mu;x), \quad |x|=1; \qquad (l=0,1,\cdots).$$

That the moment problem (179) possesses, for any given unit vector x, such

‡ See "Ha", p. 226. Cf. also I. J. Schoenberg, "On Finite and Infinite Completely Monotonic Sequences," *Bulletin of the American Mathematical Society*, Vol. 38, No. 2 (1932), pp. 72-76.

a monotone solution follows from the above lemma inasmuch as the Hausdorff conditions (175) are, for the moment problem (179)

$$\Delta^{n-m}\Phi[K^m;x] \geqq 0, \quad \text{where} \quad \Delta^{n-m}\Phi[K^m;x]$$
$$\equiv \Phi[(K^m - \binom{n-m}{1}K^{m+1} + \cdots + (-1)^{n-m}K^n);x] = \Phi[K^m(E-K)^{n-m};x],$$

and which are, from (172) and (173), obviously fulfilled. In place of the normalizing conditions (176) prescribed for the solution $\rho(\mu)$ of (174) we now prescribe

$$(180) \quad 0 \equiv \rho(0;x) \quad \text{and} \quad \rho(\mu-0;x) \equiv \rho(\mu;x) \quad \text{for} \quad 0<\mu<1; \ |x|=1,$$

for the solution $\rho(\mu;x)$ of (179) obtaining in place of (177)

$$(181) \qquad \rho(1;x) - \rho(0;x) \equiv \rho(1;x) \equiv 1 \quad \text{for} \quad |x|=1,$$

the solution $\rho(\mu;x)$ of (178) being thereby uniquely determined, for any given unit vector x, at all points of the closed interval $[0,1]$.

We now define the function $\rho(\mu;x)$ also for $-\infty<\mu<0$ and $1<\mu<+\infty$ taking it identically equal to 0 and $+1$ in the two ranges respectively. With this understanding we write in place of (178)

$$(182) \qquad \Phi[K^l;x] = \int_{-\infty}^{+\infty} \mu^l d\rho(\mu;x); \qquad (l=0,1,\cdots).$$

Let H be an arbitrary bounded Hermitian matrix. Since H is bounded there exist two real numbers α $(\alpha>0)$ and β so that the matrix

$$(183) \qquad K = \alpha H + \beta E$$

fulfills the condition (164). In denoting by $\rho(\mu;x)$ the solution of the moment problem (179), (180), (181) we introduce a function $\sigma(\mu;x)$ by means of the definition

$$(184) \quad \sigma(\mu;x) \equiv \rho(\alpha\mu+\beta;x), \quad \alpha>0, \quad -\infty<\mu<+\infty.$$

From (182), (183) and (184) we then have

$$\Phi[H^l;x] = \int_{-\infty}^{+\infty} \mu^l d\sigma(\mu;x),$$

i. e. the spectral representation of Hilbert.

A NOTE ON THE CONVERGENCE OF THE SUCCESSIVE APPROXIMATIONS TO THE SOLUTION OF AN ORDINARY DIFFERENTIAL EQUATION.

By E. K. HAVILAND.

If $f(x, y)$ is continuous in x and y together in $|x - x_0| \leqq a$, $|y - y_0| \leqq b$, where we may assume without restriction that $x_0 = y_0 = 0$, Peano has proved * the existence of at least one solution of the differential equation

$$dy/dx = f(x, y)$$

satisfying the initial condition $y(0) = 0$.

Although there will in general be more than one solution under these conditions, O. Perron,† by a slight modification of the proof of M. Nagumo,‡ has shown that a sufficient condition for the uniqueness of the solution is

$$|f(x, y') - f(x, y'')| \leqq (|k|/x)\,|y' - y''| \tag{1}$$

where

$$|k| \leqq 1, \tag{2}$$

which is seen to be a generalization of the Lipschitz condition. Furthermore, Perron has demonstrated † that no greater value of $|k|$ is sufficient.

Correspondingly, the continuity condition assumed in Peano's proof is not in general sufficient to insure the convergence of the successive approximations to the solution given by

$$y_n(x) = \int_0^x f(t, y_{n-1}(t))\,dt, \qquad y_0(x) \equiv 0$$

M. Müller being the first to give § an example where the successive approximations diverge although $f(x, y)$ is continuous.

* See, for example, E. Kamke, *Differentialgleichungen* (Leipzig, 1930), pp. 59-66.

† O. Perron, "Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnung," *Mathematische Zeitschrift*, Vol. 28 (1928), pp. 216-219.

‡ M. Nagumo, "Eine hinreichende Bedingung für die Unität der Lösung von Differentialgleichungen erster Ordnung," *Japanese Journal of Mathematics*, Vol. 3 (1926), pp. 107-112.

§ M. Müller, "Über das Fundamentaltheorem in der Theorie der gewöhnlichen Differentialgleichungen," *Mathematische Zeitschrift*, Vol. 26 (1927), pp. 619-645, especially p. 629.

A Rosenblatt has shown,* however, that (1) is a sufficient condition for the convergence of the successive approximations if condition (2) is replaced by $|k|<1$. It is accordingly of interest to consider the possibility of the occurrence of larger values of $|k|$ in a sufficient condition for the convergence of the successive approximations. In this note, it is proposed to show by a refinement of Müller's method that condition (1) is not sufficient for convergence when $|k|>1$. As similar methods have failed to establish either the sufficiency or the non-sufficiency of (1) if $|k|=1$, this boundary case remains unsolved.

The non-sufficiency of (1) for the convergence of the successive approximations when $|k|>1$ is established by the following example:

$$\text{Let } dy/dx=f(x,y), \text{ where } f(x,y)= \begin{cases} 0 & x=0, & -\infty<y<\infty \\ x^{\epsilon} & 0<x\leqq a, & -\infty<y<0 \\ x^{\epsilon}-(1+\epsilon)y/x, & 0<x\leqq a, & 0\leqq y\leqq |x|^{1+\epsilon} \\ -\epsilon x^{\epsilon} & 0<x\leqq a, & |x|^{1+\epsilon}<y<\infty \end{cases}$$

and let $x_0=y_0=0$.

Then

$$y_1(x)=\int_0^x f(t,y_0)\,dt=\int_0^x f(t,0)\,dt=\int_0^x t^{\epsilon}dt=x^{1+\epsilon}/(1+\epsilon)$$

$$y_2(x)=\int_0^x f(t,y_1(t))\,dt=\int_0^x f(t,t^{1+\epsilon}/(1+\epsilon))\,dt=0$$

and so in general

$$\left.\begin{aligned} y_{2k-1}(x)&=x^{1+\epsilon}/(1+\epsilon) \\ y_{2k}(x)&=0 \end{aligned}\right\} \qquad (k=1,2,3,\cdots).$$

Thus the successive approximations diverge. Furthermore, the limits of the convergent sequences $\{y_{2k-1}(x)\}$, $\{y_{2k}(x)\}$ do not satisfy the differential equation.

Whether or not (1) is a sufficient condition for the convergence of the successive approximations when $|k|=1$ is a question we are not at present able to answer. In the example $dy/dx=f(x,y)$, where

$$f(x,y)= \begin{cases} 0, & x=0, & -\infty<y<\infty \\ \frac{1}{2}x^{\epsilon}, & 0<x\leqq a, & -\infty<y<0 \\ \frac{1}{2}x^{\epsilon}-y/x, & 0<x\leqq a, & 0\leqq y\leqq |x|^{1+\epsilon} \\ -\frac{1}{2}x^{\epsilon}, & 0<x\leqq a, & |x|^{1+\epsilon}<y<\infty \end{cases}$$

and $x_0=y_0=0$,

---

* A. Rosenblatt, "Über die Existenz von Integralen gewöhnlicher Differentialgleichungen," *Arkiv för Matematik, Astronomi och Fysik*, Vol. 5, No. 2 (1909), pp. 1-4.

the successive approximations may be shown to converge to the solution

$$y = \tfrac{1}{2}\,(1/(2+\epsilon))x^{1+\epsilon}$$

however small the value of ϵ. In the limiting case, where $\epsilon = 0$, the successive approximations do, indeed, diverge, but the corresponding function $f(x, y)$, although bounded, is not continuous at $(0, 0)$.

Again, the functions

$$f(x,y) = \begin{cases} 0, & x = 0, & -\infty < y < \infty \\ \log(1 \pm x), & 0 < x \leqq a, & -\infty < y < 0 \\ \log(1 \pm x) - y/x, & 0 < x \leqq a, & 0 \leqq y \leqq |x| \log(1 \pm |x|) \\ 0, & 0 < x \leqq a, & |x| \log(1 \pm |x|) < y < \infty \end{cases}$$

lead to converging approximations.

In consequence, there does not appear to be any possibility of setting up a divergent sequence of successive approximations in the case $|k| = 1$ by the method of Müller. It may, of course, be that any search for a function $f(x, y)$ for which the successive approximations diverge under the condition (1) when $|k| = 1$ is futile, for this condition may be sufficient for convergence, though such has not so far been proved to be the case to the best of our knowledge.

Examples such as $dy/dx = y^{1/3}$ show that the convergence of the successive approximations is not sufficient to insure uniqueness of the solution. If an example could be given where the successive approximations diverged when $f(x, y)$ satisfied condition (1) with $|k| = 1$, it would be shown by virtue of the Nagumo-Perron theorem that conversely uniqueness was not a sufficient condition for the convergence of the successive approximations, a question at present undetermined.

THE JOHNS HOPKINS UNIVERSITY.